c/o PAULA

Access DB# 95823

SEARCH REQUEST FORM

Scientific and Technical Information Center

#3

Requester's Full Name: *Maury Audet* Examiner #: *79808* **Date:** 6/4/03
Art Unit: *1654* Phone Number: *305-5039* **Serial Number:** 09/909164
Mail Box & Bldg/Room Locat.: CM1-*11D13; 11D04* Results Format Preferred: *PAPER*

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: 7/21/00

***For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.**

Please search SEQ ID NO: 5-13, & 47.

Database search of the 3 SEQ's below will cover:

	1	2	3	4	5	6	7	8	9	10	11
1)	E	E	V	V	P	X	G	M	S	Y	S
2)	"	"	"	"	"	"	"	"	H	"	"
3)	"	"	"	"	"	"	"	"	D	"	"
4)	"	"	V	"	"	"	"	"	S	"	"

1.e. (SEQ 16) ~~11~~ 5

1.e. (SEQ 11)

1.e. (SEQ 12)

(SEQ 47)

*X = Nva, Val

TX MAUM

$ Pending Files →

****

STAFF USE ONLY

Searcher: Shaporé

Searcher Phone #: 308-4499

Searcher Location: _____

Date Searcher Picked Up: _____

Date Completed: 6/4/03

Searcher Prep & Review Time: _____

Clerical Prep Time: _____

Online Time: _____

Type of Search

NA Sequence (#)_____

AA Sequence (#)_____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) _____

PTO-1590 (8-01)

ALL APP.

```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 12:32:16 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1907 - 4 Jun 2003  VOL 138 ISS 23
FILE LAST UPDATED: 3 Jun 2003  (20030603/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.


=>
=>

=> d stat que l11
L6            13 SEA FILE=REGISTRY ABB=ON  PLU=ON  EEVVPXGMSYS/SQSP
L7             4 SEA FILE=REGISTRY ABB=ON  PLU=ON  EEVVPXGMHYS/SQSP
L8             5 SEA FILE=REGISTRY ABB=ON  PLU=ON  EEVVPXGMDYS/SQSP
L10           22 SEA FILE=REGISTRY ABB=ON  PLU=ON  L6 OR L7 OR L8
L11            1 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L10


=>
=>

=> d ibib abs hitrn l11 1

L11  ANSWER 1 OF 1  HCAPLUS  COPYRIGHT 2003 ACS
ACCESSION NUMBER:          2002:90069  HCAPLUS
DOCUMENT NUMBER:           136:145200
TITLE:                     Novel peptides as ns3-serine protease inhibitors of
                           hepatitis C virus
INVENTOR(S):               Lim-Wilby, Marguerita; Levy, Odile E.; Brunck,
                           Terrence K.
PATENT ASSIGNEE(S):        Corvas International, Inc., USA
SOURCE:                    PCT Int. Appl., 69 pp.
                           CODEN: PIXXD2
DOCUMENT TYPE:             Patent
LANGUAGE:                  English
FAMILY ACC. NUM. COUNT:    1
PATENT INFORMATION:

     PATENT NO.        KIND  DATE          APPLICATION NO.  DATE
     ----------------  ----  --------      ---------------  --------
     WO 2002008251      A2   20020131      WO 2001-US23169  20010719
     WO 2002008251      A3   20030109
         W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
             CO, CR, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, HR, HU,
             ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV, MA, MD,
             MG, MK, MN, MX, MZ, NO, NZ, PL, PT, RO, RU, SE, SG, SI, SK, SL,
```

APP.

TJ, TM, TR, TT, TZ, UA, UZ, VN, YU, ZA, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

US 2002068702	A1	20020606	US 2001-909164	20010719
EP 1301527	A2	20030416	EP 2001-955916	20010719

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR

PRIORITY APPLN. INFO.:	US 2000-220101P	P	20000721
	WO 2001-US23169	W	20010719

OTHER SOURCE(S): MARPAT 136:145200

AB The present invention discloses novel peptide compds. contg. eleven amino acid residues which have hepatitis C virus (HCV) protease inhibitory activity as well as methods for prepg. such compds. In another embodiment, the invention discloses pharmaceutical compns. comprising such peptides as well as methods of using them to treat disorders assocd. with the HCV protease.

IT **393513-23-0DP**, MBHA-resin-bound **393513-24-1DP**, MBHA-resin-bound
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(novel peptides as ns3-serine protease inhibitors of hepatitis C virus)

IT **393512-68-0P**
RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(novel peptides as ns3-serine protease inhibitors of hepatitis C virus and combination with other antiviral agents)

IT **393512-69-1 393512-70-4 393512-71-5**
393512-72-6 393512-73-7 393512-75-9
393512-76-0 393512-77-1 393513-06-9
393513-07-0 393513-08-1 393513-09-2
393513-10-5 393513-11-6 393513-12-7
393513-13-8 393513-14-9 393513-15-0
393513-16-1
RL: PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(novel peptides as ns3-serine protease inhibitors of hepatitis C virus and combination with other antiviral agents)

=> fil caold
FILE 'CAOLD' ENTERED AT 12:32:42 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

FILE COVERS 1907-1966
FILE LAST UPDATED: 01 May 1997 (19970501/UP)

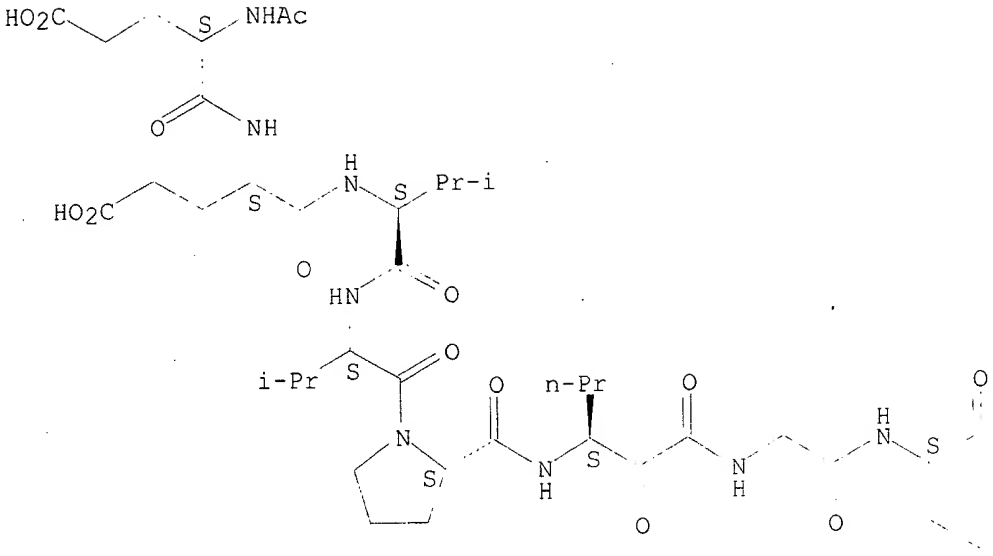
This file contains CAS Registry Numbers for easy and accurate substance identification. Title keywords, authors, patent assignees, and patent information, e.g., patent numbers, are now searchable from 1907-1966. TIFF images of CA abstracts printed between 1907-1966 are available in the PAGE display formats.

This file supports REG1stRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

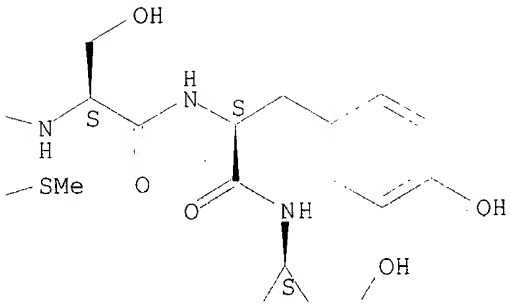
=>

```
=>

=> s l10
L12           0 L10

=>
=>

=> fil reg
FILE 'REGISTRY' ENTERED AT 12:32:49 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES:     3 JUN 2003  HIGHEST RN 524916-37-8
DICTIONARY FILE UPDATES:    3 JUN 2003  HIGHEST RN 524916-37-8

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

  Please note that search-term pricing does apply when
  conducting SmartSELECT searches.

Crossover limits have been increased.  See HELP CROSSOVER for details.

Experimental and calculated property data are now available.  See HELP
PROPERTIES for more information.  See STNote 27, Searching Properties
in the CAS Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

=>
=>

=> d sqide 1-22 l10

L10  ANSWER 1 OF 22  REGISTRY  COPYRIGHT 2003 ACS
RN   393513-24-1  REGISTRY
CN   L-Serine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-
     prolyl-(3S)-3-amino-2-oxohexanoylglycyl-L-methionyl-L-seryl-L-tyrosyl-
     (9CI)  (CA INDEX NAME)
FS   PROTEIN SEQUENCE; STEREOSEARCH
SQL  11
NTE  modified
----------------------------------------------------------------------
 type            ------ location ------         description
----------------------------------------------------------------------
terminal mod.    Glu-1           -          N-acetyl
terminal mod.    Ser-11          -          C-terminal amide
uncommon         Oaa-6           -          -
----------------------------------------------------------------------

SEQ      1 EEVVPXGMSY S
           ========== =
HITS AT:   1-11

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**
MF   C55 H83 N11 O21 S
SR   CA
LC   STN Files:   CA, CAPLUS, USPATFULL

Absolute stereochemistry.
```

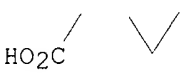
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 2 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-23-0 REGISTRY
CN L-Serine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-[[[(diphenylmethyl)amino]carbonyl]hydrazono]hexanoyl glycyl-L-methionyl-O-(1,1-dimethylethyl)-L-seryl-O-(1,1-dimethylethyl)-L-tyrosyl-O-(1,1-dimethylethyl)-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified (modifications unspecified)

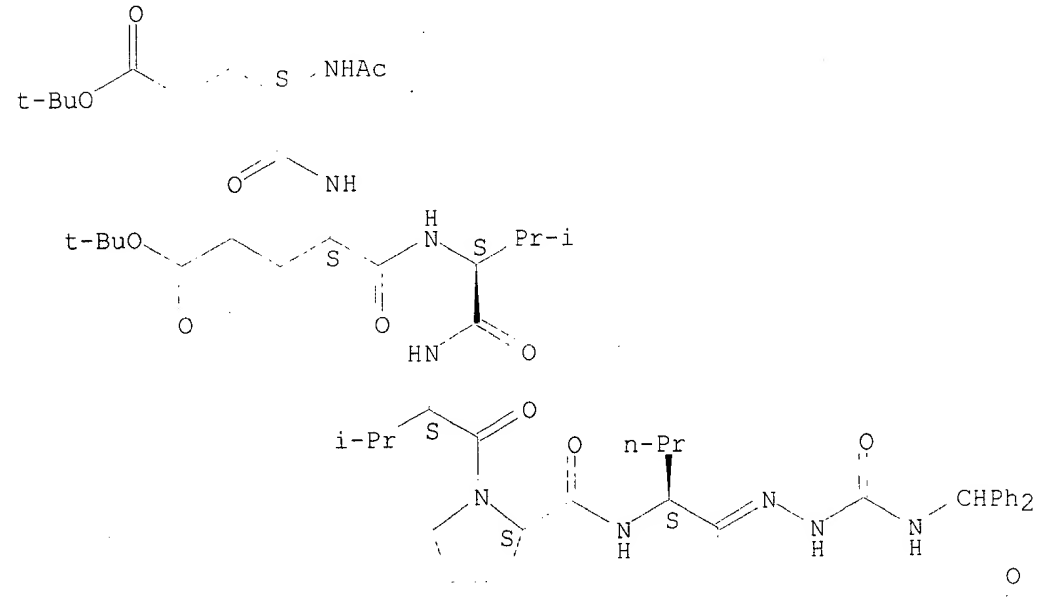
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

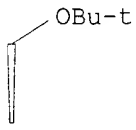
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C89 H136 N14 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.
Double bond geometry unknown.

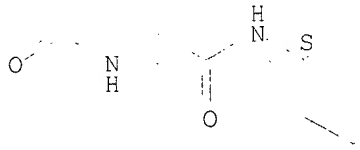
PAGE 1-A



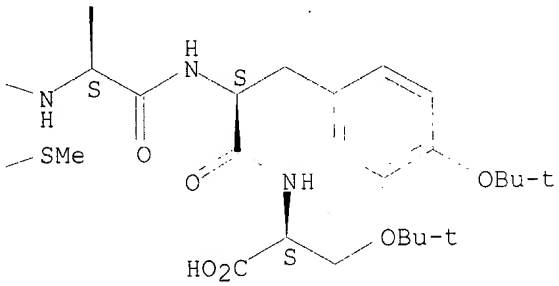
PAGE 1-B



PAGE 2-A



PAGE 2-B



1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 3 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-16-1 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-3-amino-2-oxo-5-hexynoylglycyl-L-methionyl-L-seryl-L-

tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

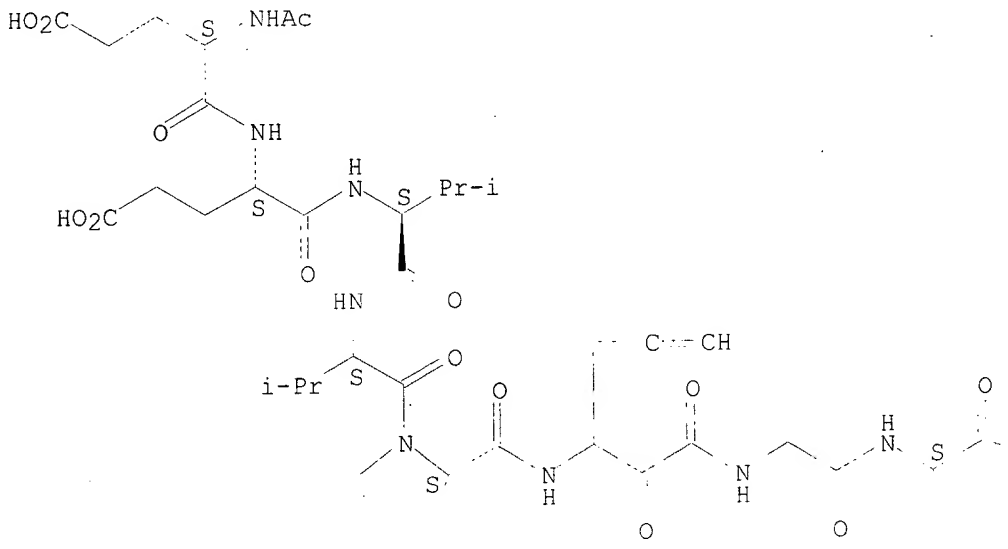
type	location		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

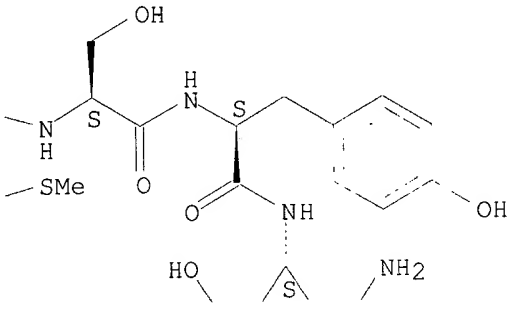
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H80 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

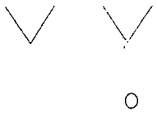
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 4 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-15-0 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S,4S)-3-amino-4-hydroxy-2-oxopentanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

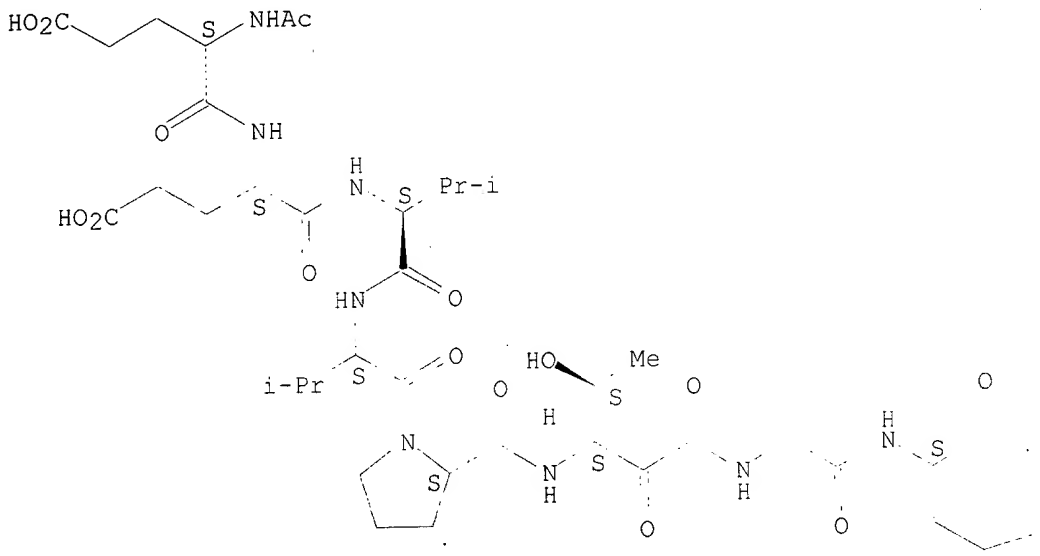
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

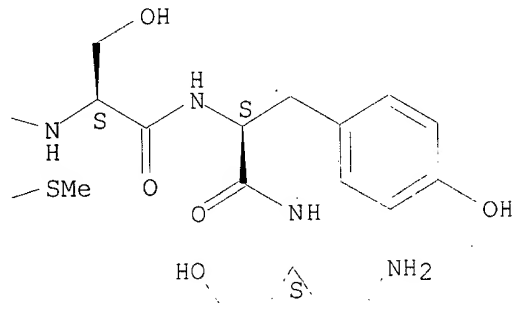
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C54 H82 N12 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

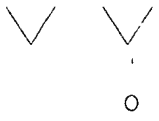
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 5 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-14-9 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxopentanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

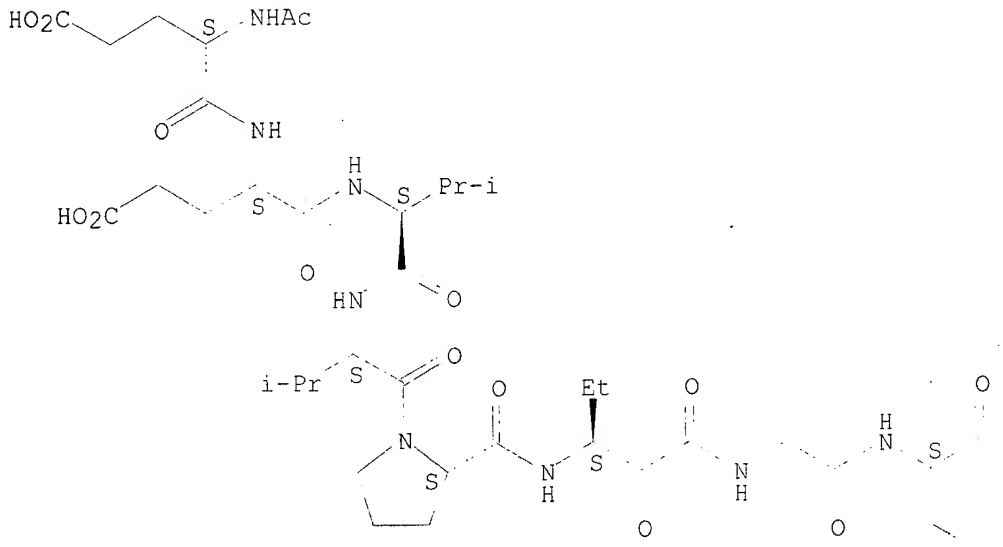
type	location		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

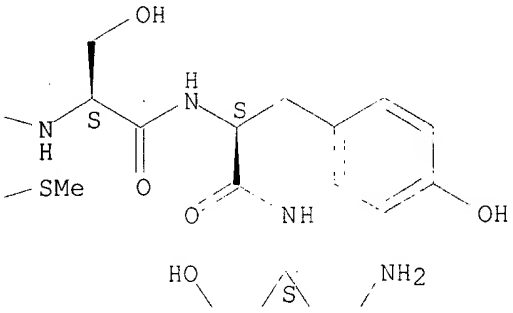
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C54 H82 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

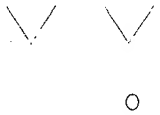
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 6 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-13-8 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxoheptanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

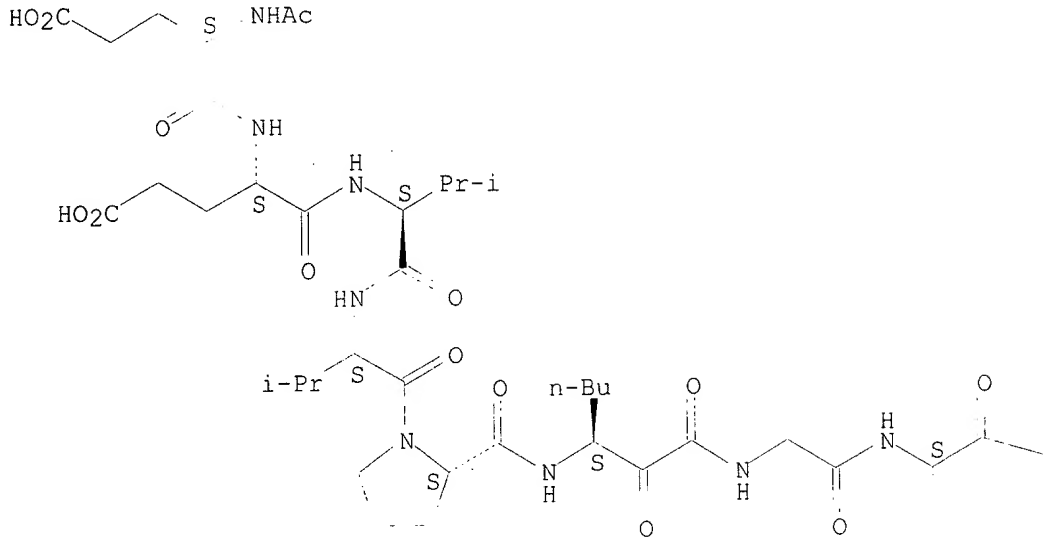
SEQ 1 EEVVPXGMSY S
========== =

HITS AT: 1-11

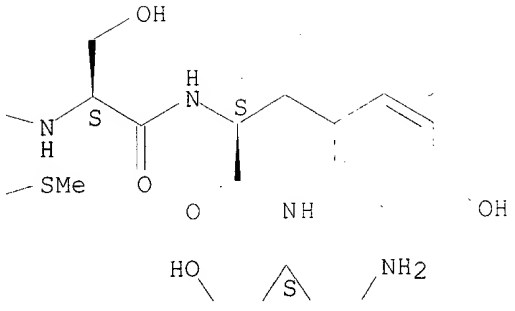
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H86 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

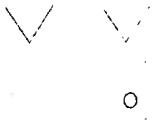
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

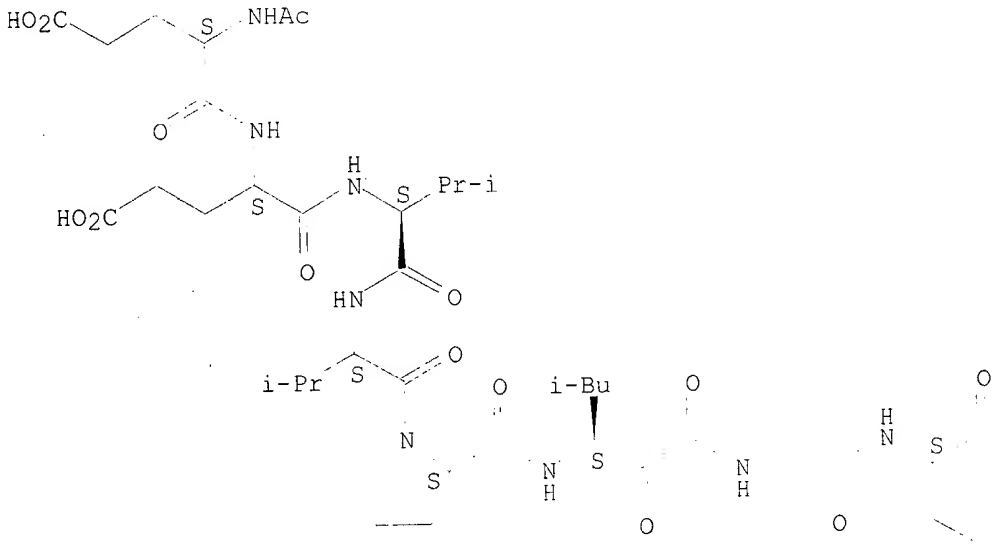
L10 ANSWER 7 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-12-7 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-5-methyl-2-oxohexanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

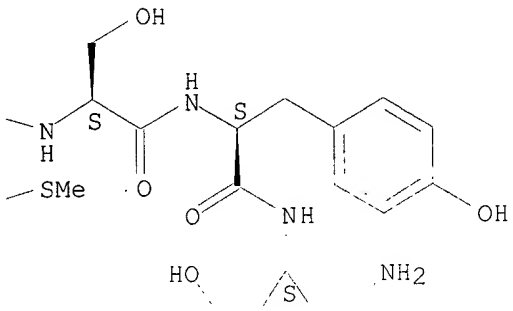
SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H86 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

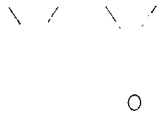
Absolute stereochemistry.



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 8 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-11-6 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-4-methyl-2-oxopentanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

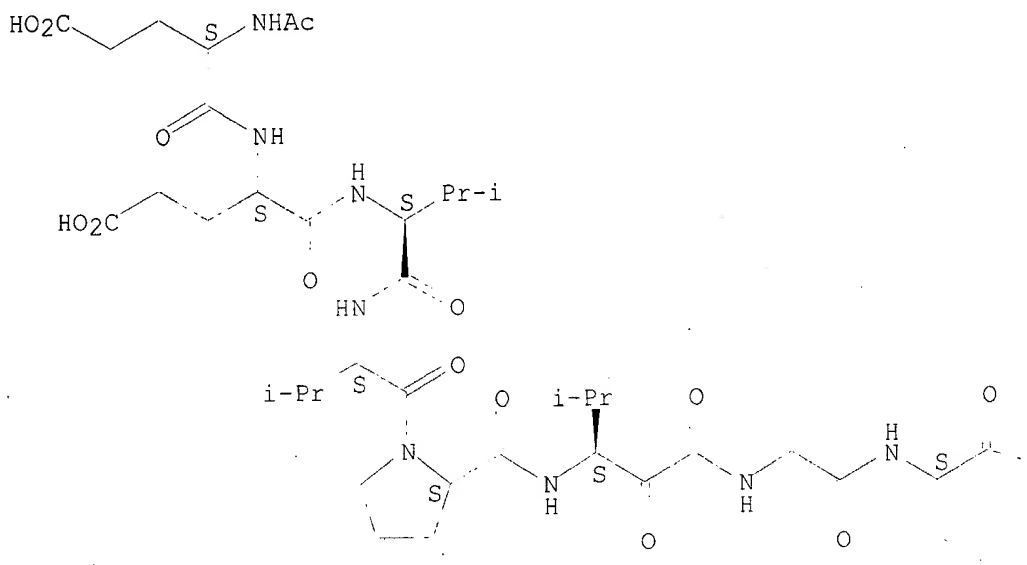
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

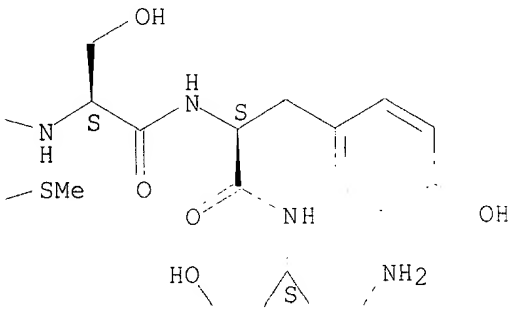
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

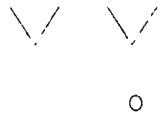
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 9 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-10-5 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-(2R)-2-amino-4-(methylsulfinyl)butanoyl-D-.alpha.-aspartyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

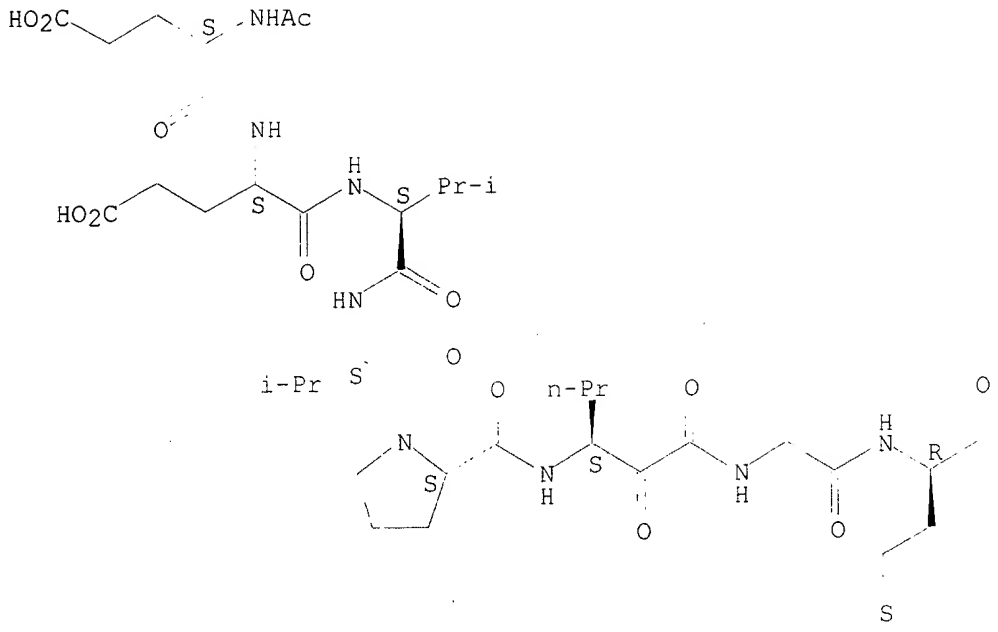
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D
stereo	Asp-9	-	D

SEQ 1 EEVVPXGMDY S
========== =
HITS AT: 1-11

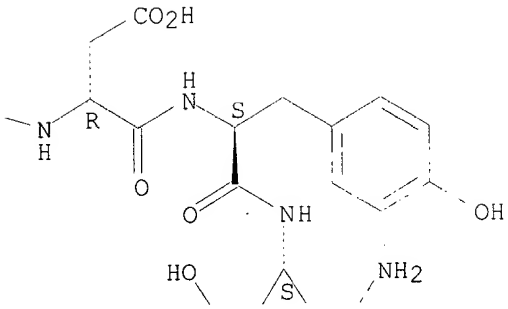
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H84 N12 O22 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A

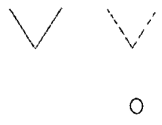


PAGE 1-B



PAGE 2-A

Me O



O

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 10 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-09-2 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-(2R)-2-amino-4-(methylsulfinyl)butanoyl-L-.alpha.-aspartyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

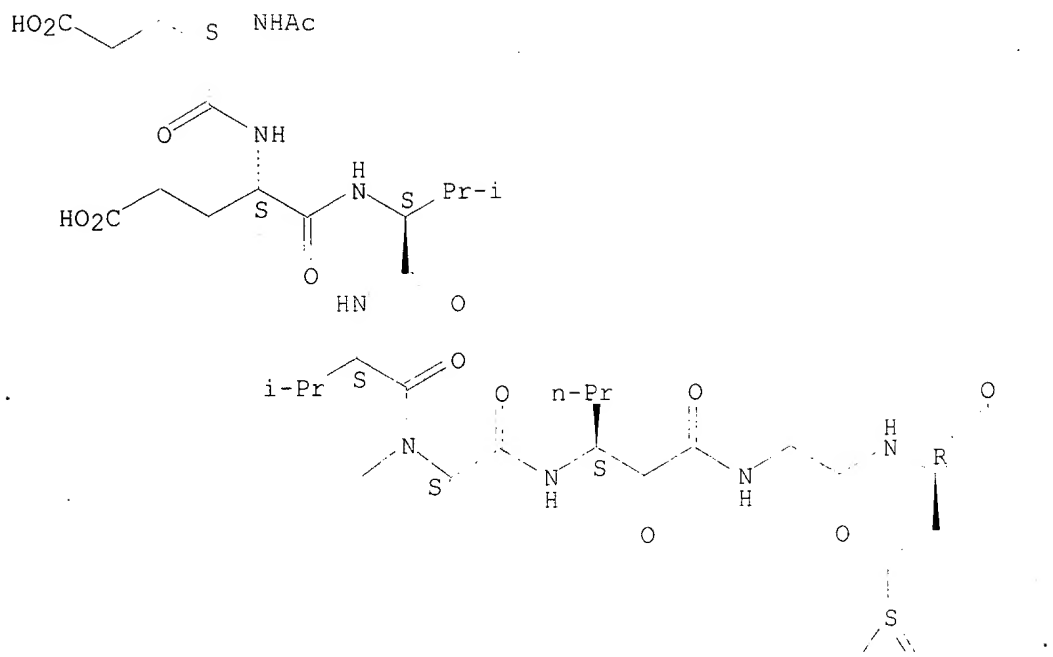
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D

SEQ 1 EEVVPXGMDY S
========== =
HITS AT: 1-11

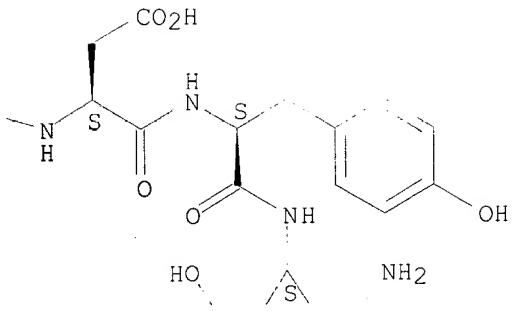
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H84 N12 O22 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

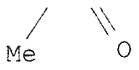
PAGE 1-A



PAGE 1-B



PAGE 2-A



O

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 11 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-08-1 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-(2R)-2-amino-4-(methylsulfinyl)butanoyl-D-histidyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

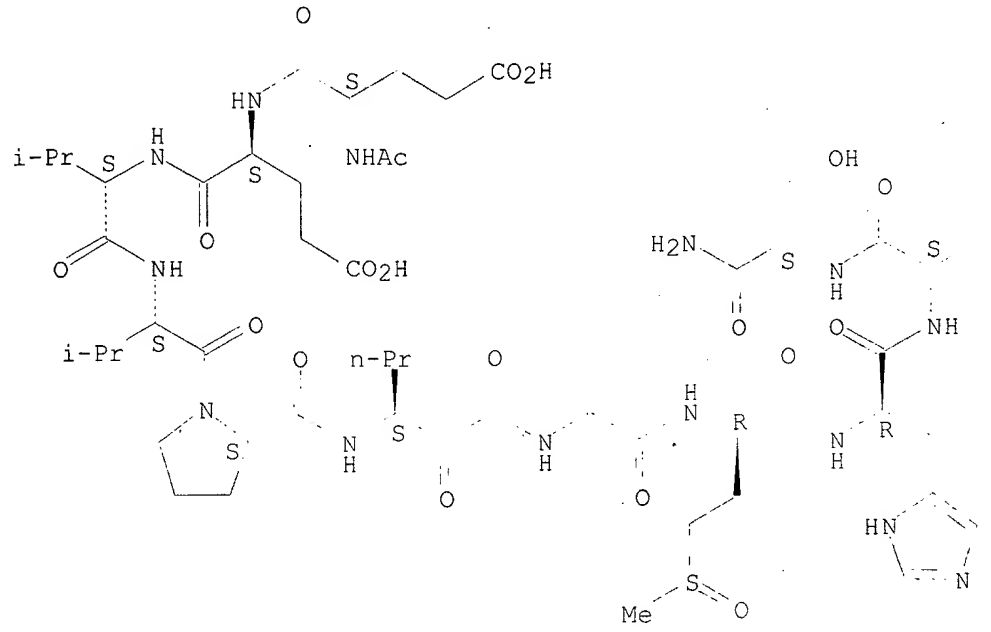
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D
stereo	His-9	-	D

SEQ 1 EEVVPXGMHY S
========== =
HITS AT: 1-11

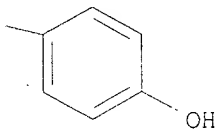
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C58 H86 N14 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 12 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-07-0 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-(2R)-2-amino-4-(methylsulfinyl)butanoyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

type	------ location ------	description

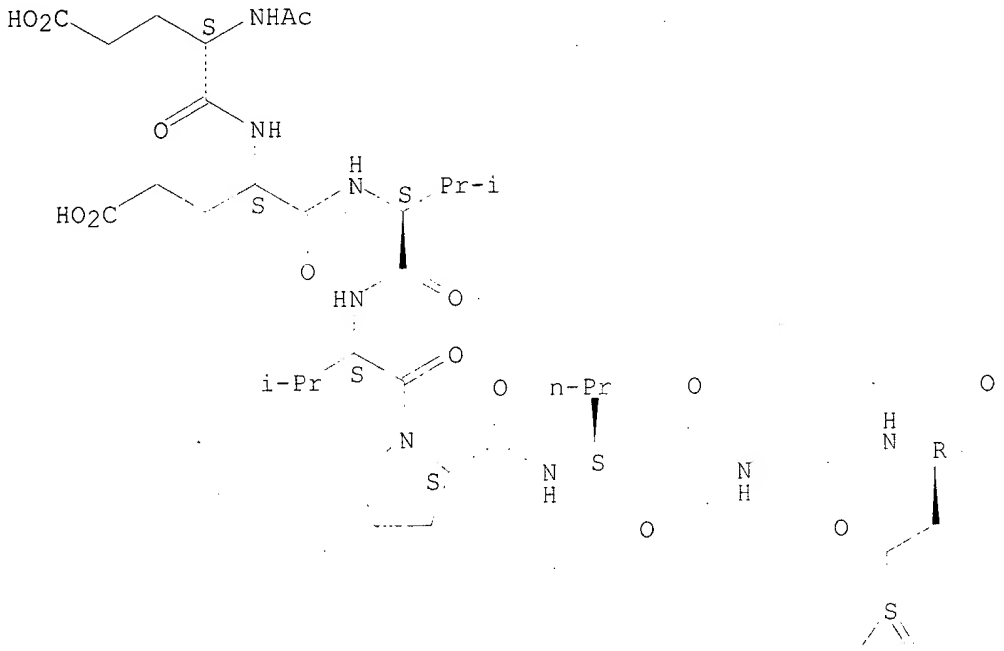
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

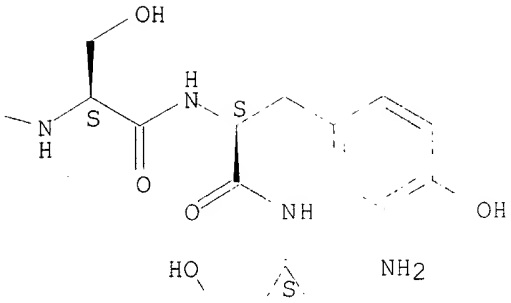
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PAGE 2-A

Me O

PAGE 2-B

O

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 13 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393513-06-9 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-(2S)-2-amino-4-(methylsulfinyl)butanoyl-L-histidyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide

uncommon Oaa-6 - -

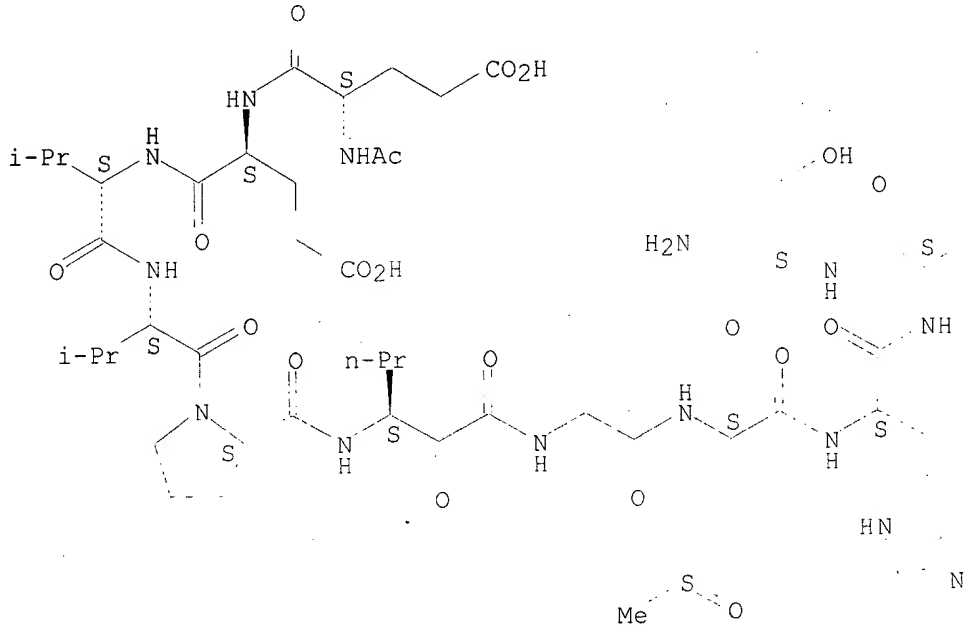
SEQ 1 EEVVPXGMHY S
========== =

HITS AT: 1-11

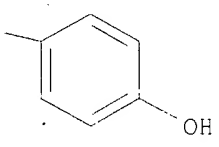
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C58 H86 N14 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 14 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-77-1 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-D-methionyl-D-.alpha.-aspartyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

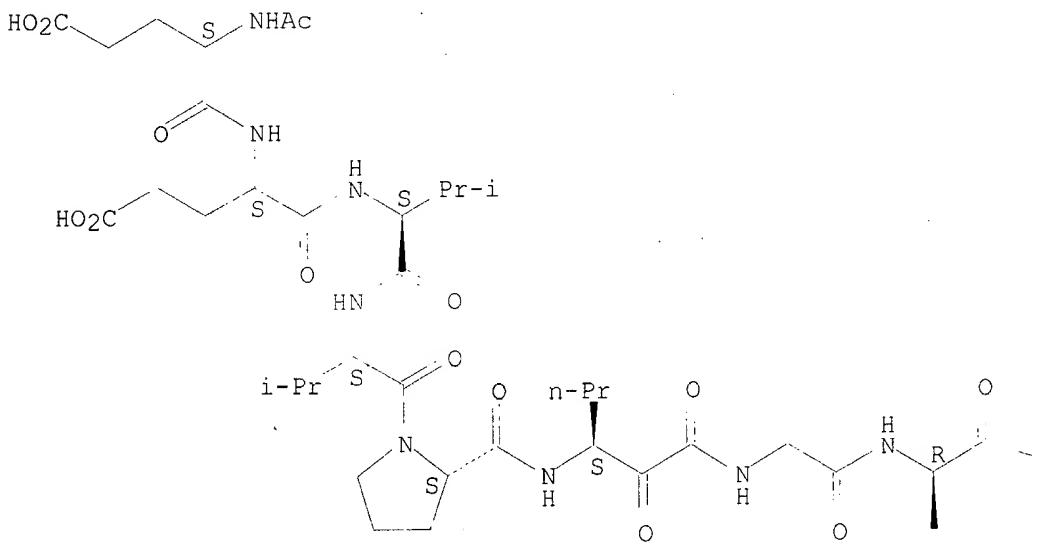
type	location		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D
stereo	Asp-9	-	D

SEQ 1 EEVVPXGMDY S
========== =
HITS AT: 1-11

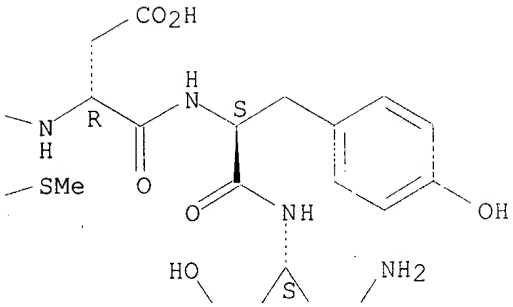
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H84 N12 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

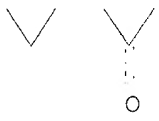
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 15 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-76-0 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-D-methionyl-L-.alpha.-aspartyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

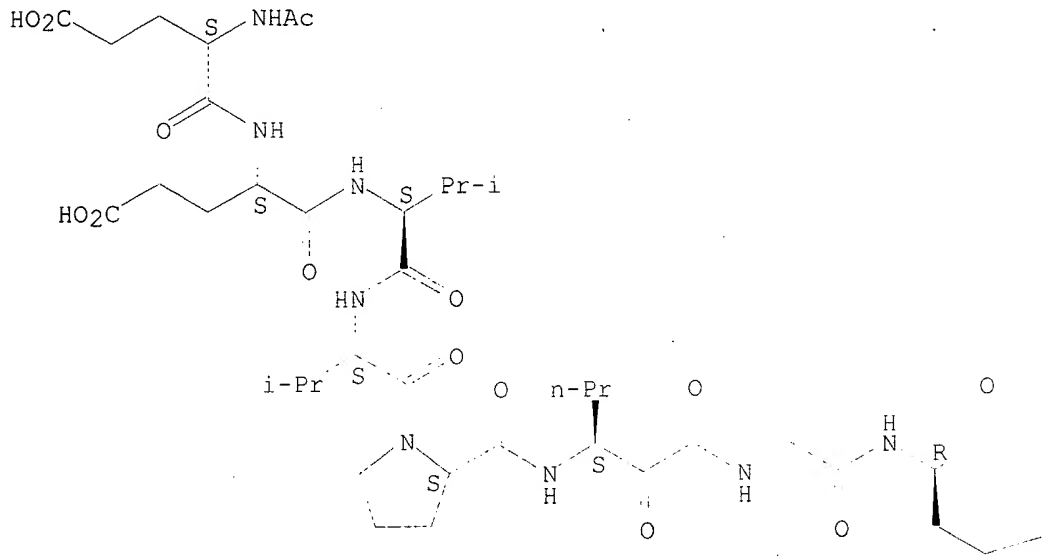
type	location		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D

SEQ 1 EEVVPXGMDY S
========== =
HITS AT: 1-11

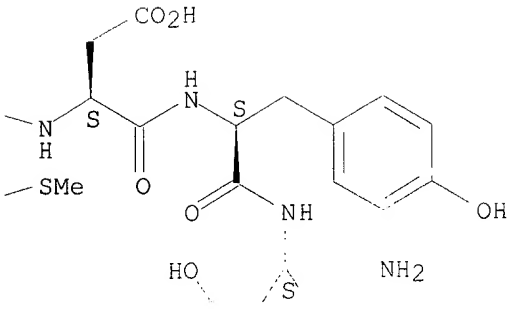
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H84 N12 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

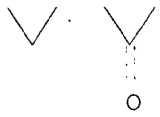
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 16 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-75-9 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-D-methionyl-L-histidyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

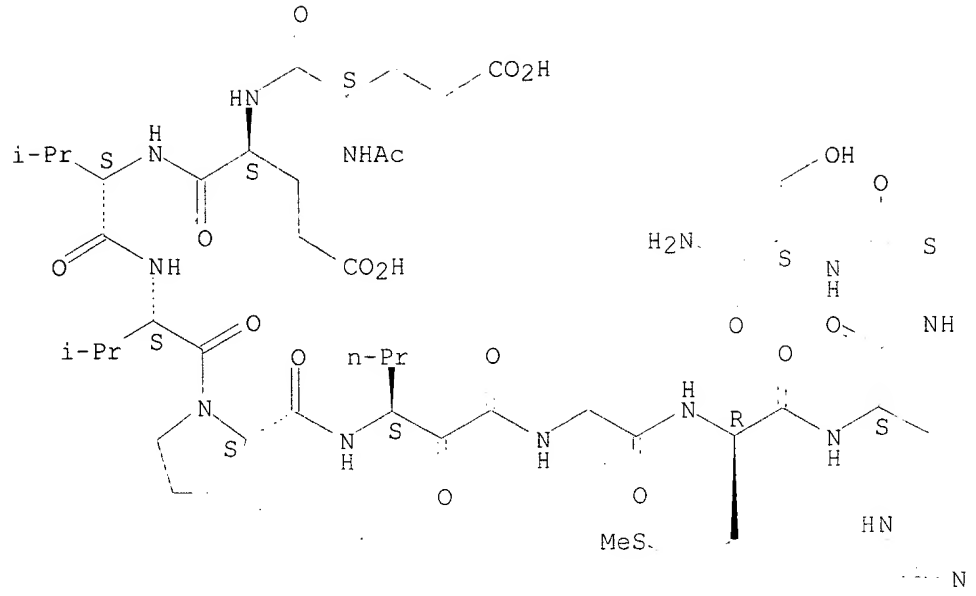
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D

SEQ 1 EEVVPXGMHY S
========== =
HITS AT: 1-11

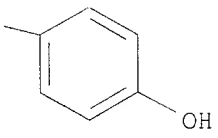
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C58 H86 N14 O19 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 17 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-73-7 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-D-methionyl-D-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

type	------ location ------	description

terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D
stereo	Ser-9	-	D

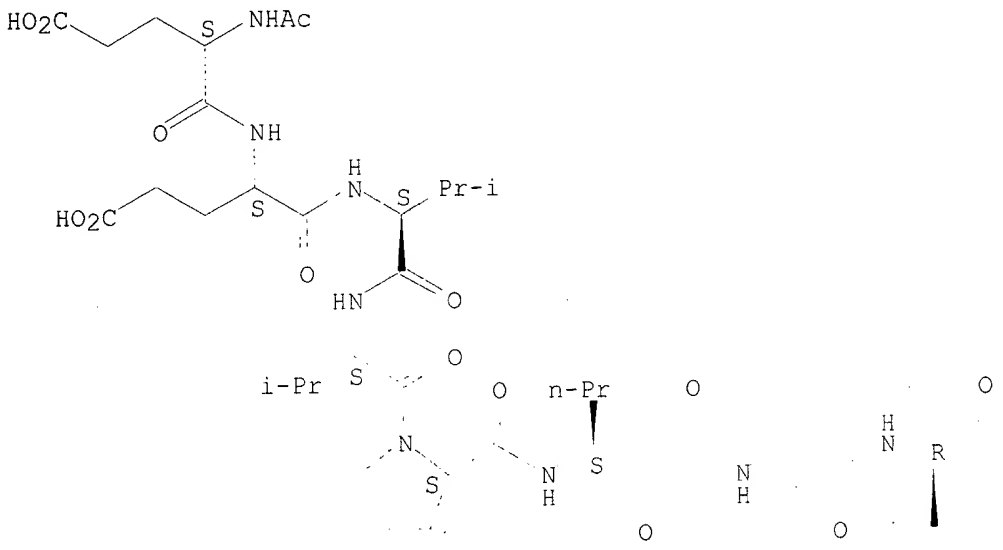
SEQ 1 EEVVPXGMSY S
========== =

HITS AT: 1-11

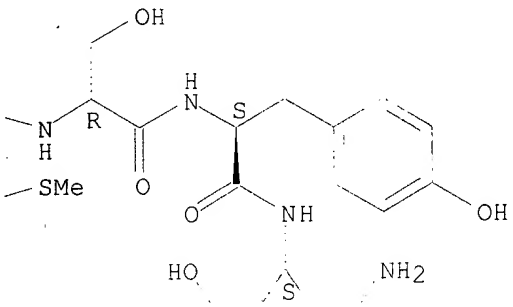
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

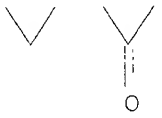
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 18 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-72-6 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-D-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

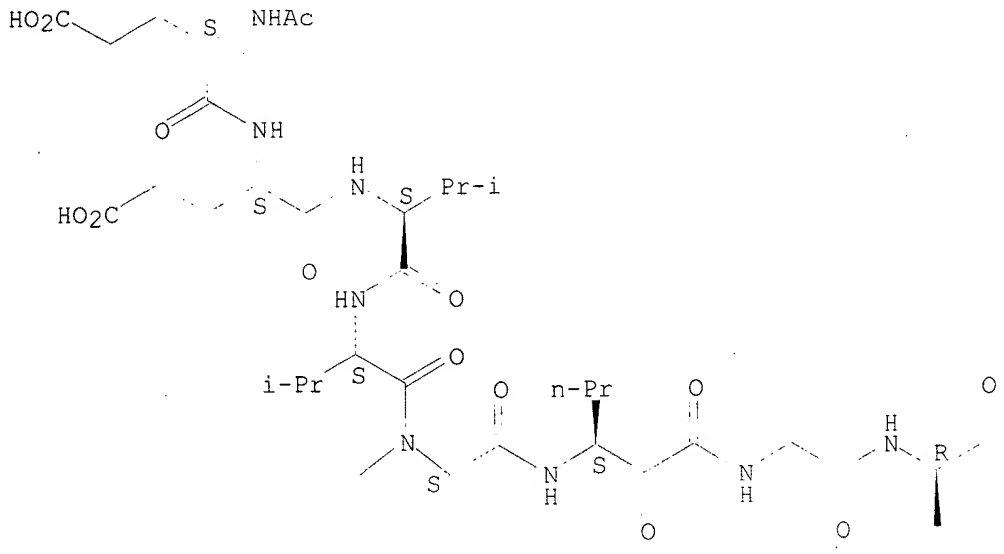
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Met-8	-	D

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

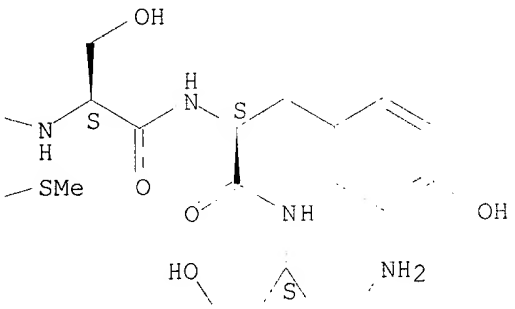
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

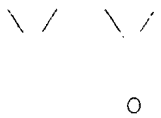
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 19 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-71-5 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-L-methionyl-D-.alpha.-aspartyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

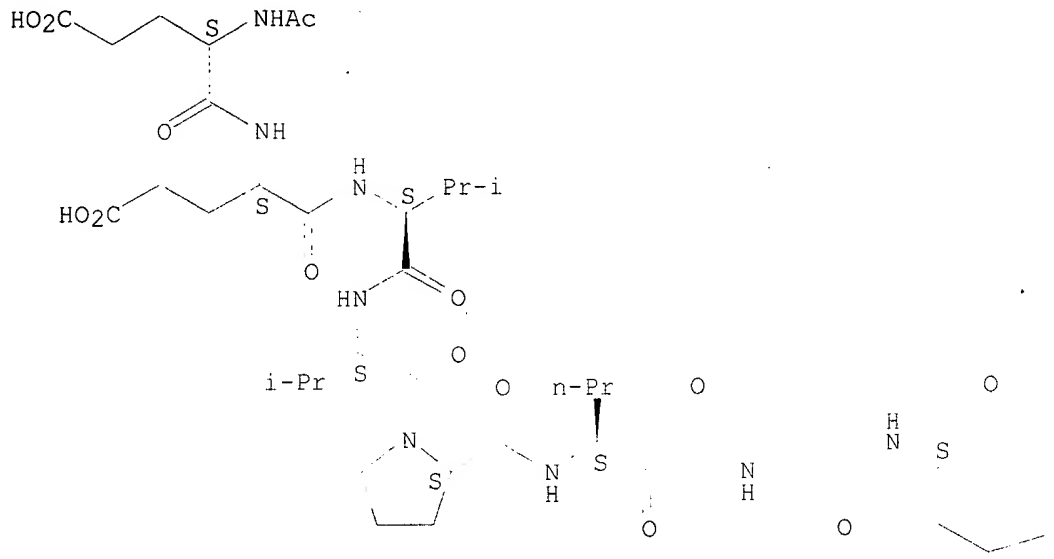
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Asp-9	-	D

SEQ 1 EEVVPXGMDY S
========== =
HITS AT: 1-11

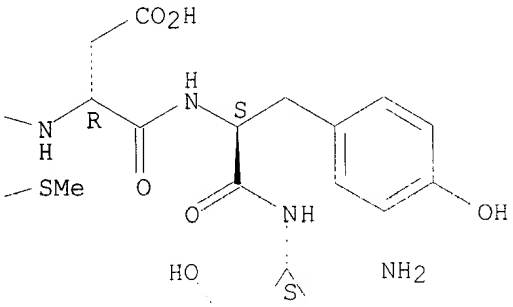
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C56 H84 N12 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

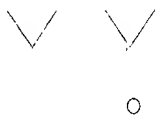
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 20 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-70-4 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-L-methionyl-D-histidyl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

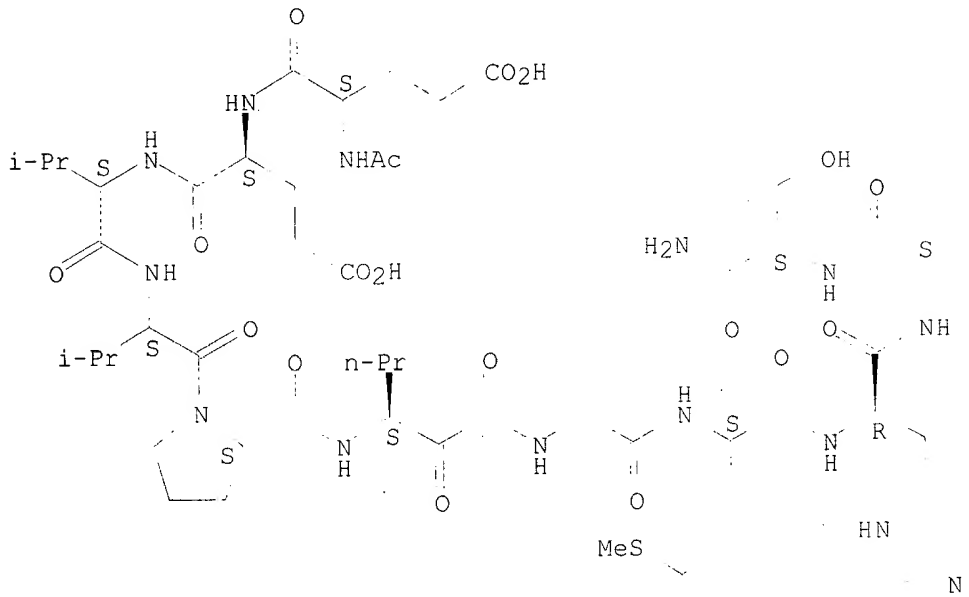
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	His-9	-	D

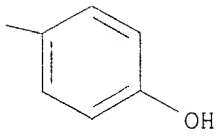
SEQ 1 EEVVPXGMHY S
========== =
HITS AT: 1-11

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C58 H86 N14 O19 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 21 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-69-1 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-L-methionyl-D-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

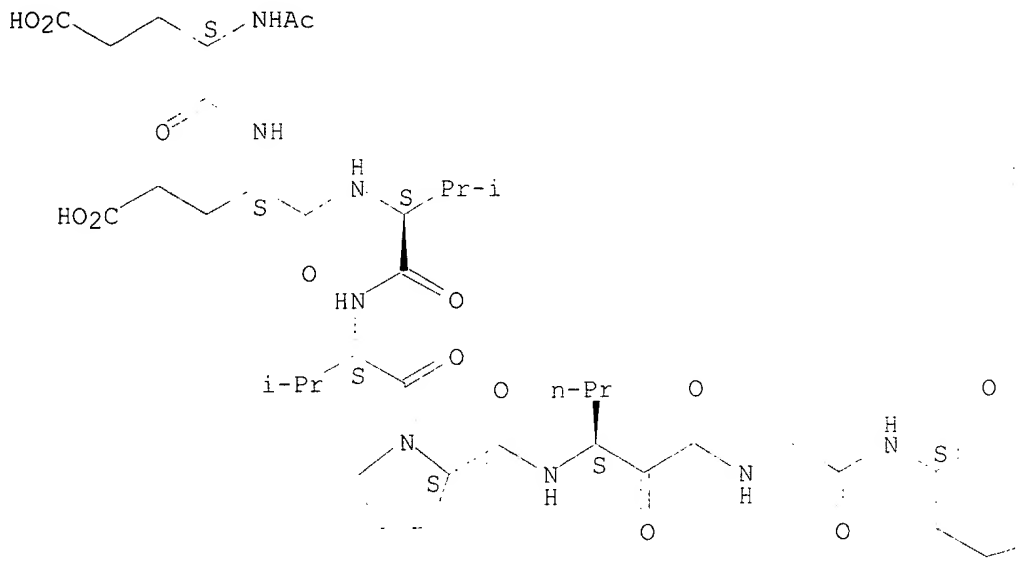
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-
stereo	Ser-9	-	D

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

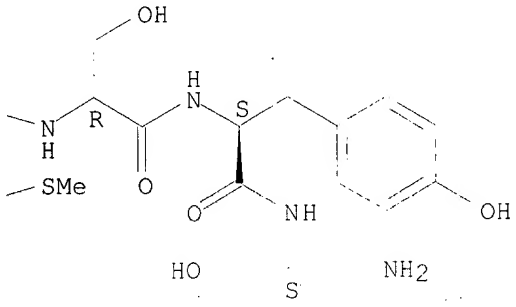
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

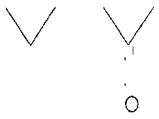
PAGE 1-A



PAGE 1-B



PAGE 2-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L10 ANSWER 22 OF 22 REGISTRY COPYRIGHT 2003 ACS
RN 393512-68-0 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoylglycyl-L-methionyl-L-seryl-L-tyrosyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified

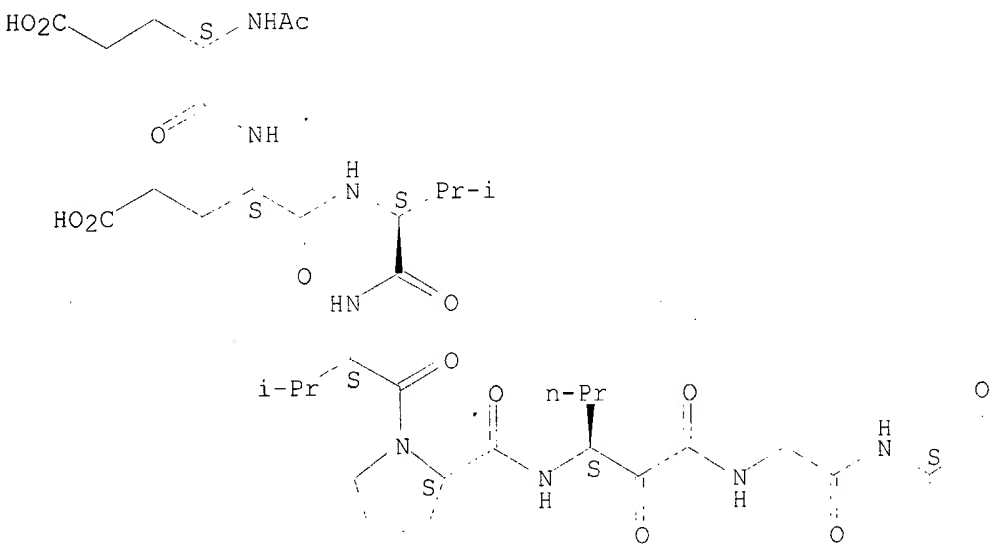
type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
terminal mod.	Ser-11	-	C-terminal amide
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S
========== =
HITS AT: 1-11

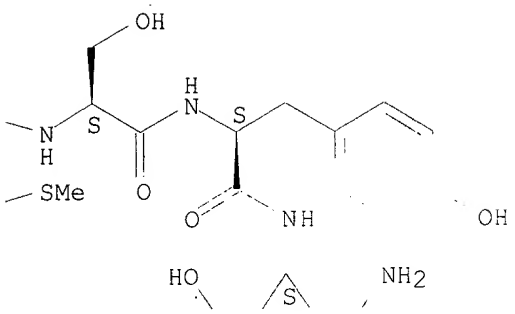
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C55 H84 N12 O20 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PAGE 2-B

O

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

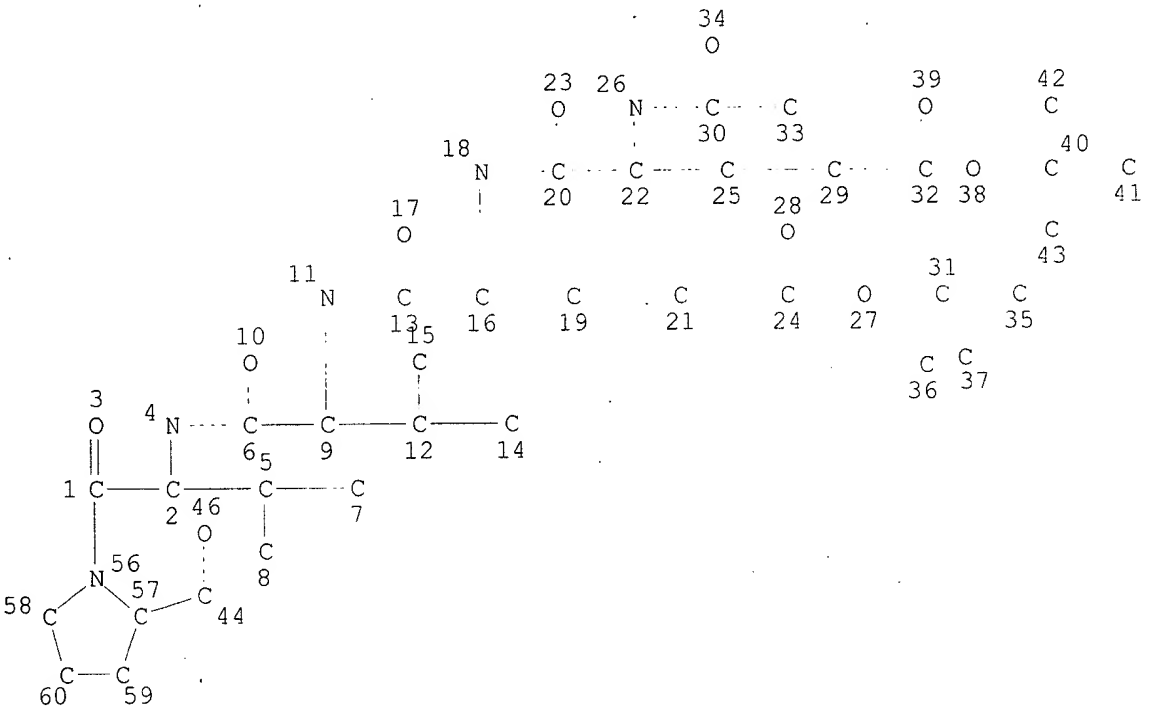
=>
=>

=> d stat que l9
L9 0 SEA FILE=REGISTRY ABB=ON PLU=ON EEVVPVGMSYS/SQSP

```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 15:10:35 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1907 - 4 Jun 2003  VOL 138 ISS 23
FILE LAST UPDATED: 3 Jun 2003  (20030603/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.


=>
=>

=> d stat que
L1            16 SEA FILE=REGISTRY ABB=ON  PLU=ON  393520-23-5/RN OR 393520-17-7
                 /RN OR 393520-15-5/RN OR 393520-11-1/RN OR 393520-09-7/RN OR
                 393525-39-8/RN OR 393520-00-8/RN OR 393524-07-7/RN OR 393521-98
                 -7/RN OR 393521-84-1/RN OR 393520-07-5/RN OR 393520-02-0/RN OR
                 393521-82-9/RN OR 393521-33-0/RN OR 393521-31-8/RN OR 393520-97
                 -3/RN
L2             3 SEA FILE=REGISTRY ABB=ON  PLU=ON  393520-29-1/RN OR 393520-27-9
                 /RN OR 393520-25-7/RN
L3            19 SEA FILE=REGISTRY ABB=ON  PLU=ON  L1 OR L2
L4                STR
```



NODE ATTRIBUTES:
DEFAULT MLEVEL IS ATOM
DEFAULT ECLEVEL IS LIMITED

GRAPH ATTRIBUTES:
RING(S) ARE ISOLATED OR EMBEDDED
NUMBER OF NODES IS 50

STEREO ATTRIBUTES: NONE
L6 32 SEA FILE=REGISTRY SSS FUL L4
L7 24 SEA FILE=REGISTRY ABB=ON PLU=ON L6 NOT L3
L8 2 SEA FILE=HCAPLUS ABB=ON PLU=ON L7

=>
=>

=> d ibib abs hitrn 18 1-2

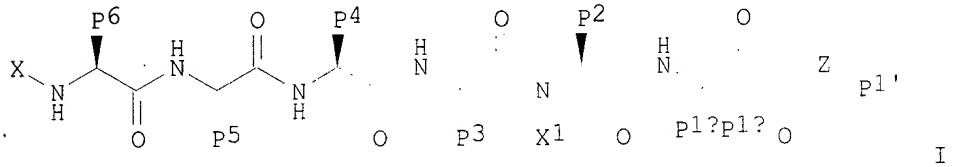
L8 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:90074 HCAPLUS
DOCUMENT NUMBER: 136:151440
TITLE: Preparation of novel peptides as NS3-serine protease inhibitors of hepatitis C virus
INVENTOR(S): Saksena, Anil K.; Girijavallabhan, Viyyoor Moopil; Lovey, Raymond G.; Jao, Edwin E.; Bennett, Frank; McCormick, Jinping; Wang, Haiyan; Pike, Russell E.; Bogen, Stephane L.; Liu, Yi-Tsung; Arasappan, Ashok; Parekh, Tejal; Pinto, Patrick A.; Njoroge, F. George; Ganguly, Ashit K.; Brunck, Terence K.; Kemp, Scott Jeffrey; Levy, Odile Esther; Lim-Wilby, Marguerita
PATENT ASSIGNEE(S): Schering Corporation, USA; Corvas International, Inc.
SOURCE: PCT Int. Appl., 197 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent

LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

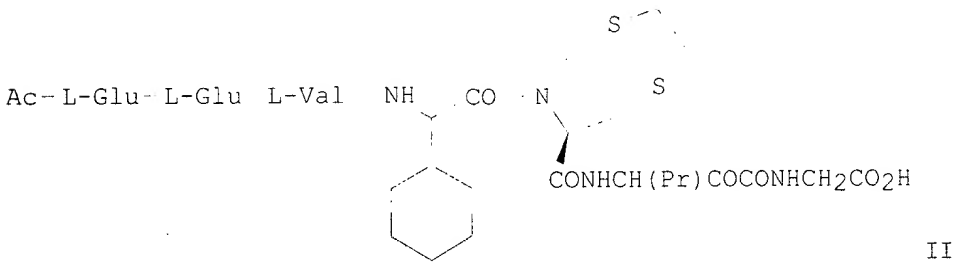
PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002008256	A2	20020131	WO 2001-US22826	20010719
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, HR, HU, ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV, MA, MD, MG, MK, MN, MX, MZ, NO, NZ, PL, PT, RO, RU, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UZ, VN, YU, ZA, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
US 2003036501	A1	20030220	US 2001-909062	20010719
EP 1301528	A2	20030416	EP 2001-959046	20010719
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR				

PRIORITY APPLN. INFO.: US 2000-220109P P 20000721
WO 2001-US22826 W 20010719

OTHER SOURCE(S): MARPAT 136:151440
GI



I



II

AB Novel peptides I [Z = O, NH or substituted imino; X = (un)substituted alkylsulfonyl, heterocyclylsulfonyl, heterocyclylalkylsulfonyl, arylsulfonyl, heteroarylsulfonyl, alkylcarbonyl, heterocyclylcarbonyl, heterocyclylalkylcarbonyl, arylcarbonyl, heteroarylcarbonyl, alkoxycarbonyl, heterocyclyloxycarbonyl, aryloxycarbonyl, heteroaryloxycarbonyl, alkyaminocarbonyl, heterocyclylaminocarbonyl, arylaminocarbonyl, or heteroarylaminocarbonyl; X1 = H, alkyl, arylmethyl; P1a, P1b, P2-P6 = H, (un)substituted alkyl, alkenyl, cycloalkyl, heterocyclyl, cycloalkylalkyl, heterocyclylalkyl, aryl, heteroaryl, arylalkyl, or heteroarylalkyl; P1a and P1b may optionally be joined to each other to form a spirocyclic or spiroheterocyclic ring contg. 0-6 oxygen, nitrogen, sulfur, or phosphorus atoms; P1' = H, (un)substituted alkyl, alkenyl, alkynyl, cycloalkyl, cycloalkylalkyl, heterocyclyl,

heterocyclylalkyl, aryl, arylalkyl, heteroaryl, or heteroarylalkyl] having HCV protease inhibitory activity are disclosed. Thus, peptide II was prepd. via peptide coupling in soln. and showed Ki = 1-100 nM for inhibition of HCV protease.

IT **393519-93-2P 393520-05-3P**
RL: PAC (Pharmacological activity); RCT (Reactant); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); RACT (Reactant or reagent); USES (Uses)
(prepn. of novel peptides as NS3-serine protease inhibitors of hepatitis C virus)

IT **393520-19-9P 393520-33-7P 393520-49-5P**
393520-51-9P 393520-61-1P 393520-63-3P
393520-81-5P 393520-83-7P 393520-87-1P
393521-13-6P
RL: PAC (Pharmacological activity); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(prepn. of novel peptides as NS3-serine protease inhibitors of hepatitis C virus)

IT **393513-18-3P 393525-21-8P 393525-23-0P**
393525-25-2P 393525-27-4P 393525-29-6P
393525-31-0P 393525-37-6P 393525-40-1P
393525-42-3P 393525-43-4P
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(prepn. of novel peptides as NS3-serine protease inhibitors of hepatitis C virus)

L8 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:90069 HCAPLUS
DOCUMENT NUMBER: 136:145200
TITLE: Novel peptides as ns3-serine protease inhibitors of hepatitis C virus
INVENTOR(S): Lim-Wilby, Marguerita; Levy, Odile E.; Brunck, Terrence K.
PATENT ASSIGNEE(S): Corvas International, Inc., USA
SOURCE: PCT Int. Appl., 69 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002008251	A2	20020131	WO 2001-US23169	20010719
WO 2002008251	A3	20030109		
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, HR, HU, ID, IL, IN, IS, JP, KG, KR, KZ, LC, LK, LR, LT, LU, LV, MA, MD, MG, MK, MN, MX, MZ, NO, NZ, PL, PT, RO, RU, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UZ, VN, YU, ZA, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				
US 2002068702	A1	20020606	US 2001-909164	20010719
EP 1301527	A2	20030416	EP 2001-955916	20010719
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR				

PRIORITY APPLN. INFO.: US 2000-220101P P 20000721
WO 2001-US23169 W 20010719
OTHER SOURCE(S): MARPAT 136:145200

AB The present invention discloses novel peptide compds. contg. eleven amino acid residues which have hepatitis C virus (HCV) protease inhibitory activity as well as methods for prepg. such compds. In another embodiment, the invention discloses pharmaceutical compns. comprising such peptides as well as methods of using them to treat disorders assocd. with the HCV protease.

IT **393513-18-3DP**, 2ClTrt- resin-bound **393513-18-3P**
393513-23-0DP, MBHA-resin-bound
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(novel peptides as ns3-serine protease inhibitors of hepatitis C virus)

=> fil caold
FILE 'CAOLD' ENTERED AT 15:10:53 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

FILE COVERS 1907-1966
FILE LAST UPDATED: 01 May 1997 (19970501/UP)

This file contains CAS Registry Numbers for easy and accurate substance identification. Title keywords, authors, patent assignees, and patent information, e.g., patent numbers, are now searchable from 1907-1966. TIFF images of CA abstracts printed between 1907-1966 are available in the PAGE display formats.

This file supports REG1stRY for direct browsing and searching of all substance data from the REGISTRY file. Enter HELP FIRST for more information.

=>
=>

=> s l7
L9 0 L7

=> fil reg
FILE 'REGISTRY' ENTERED AT 15:11:04 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 3 JUN 2003 HIGHEST RN 524916-37-8
DICTIONARY FILE UPDATES: 3 JUN 2003 HIGHEST RN 524916-37-8

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

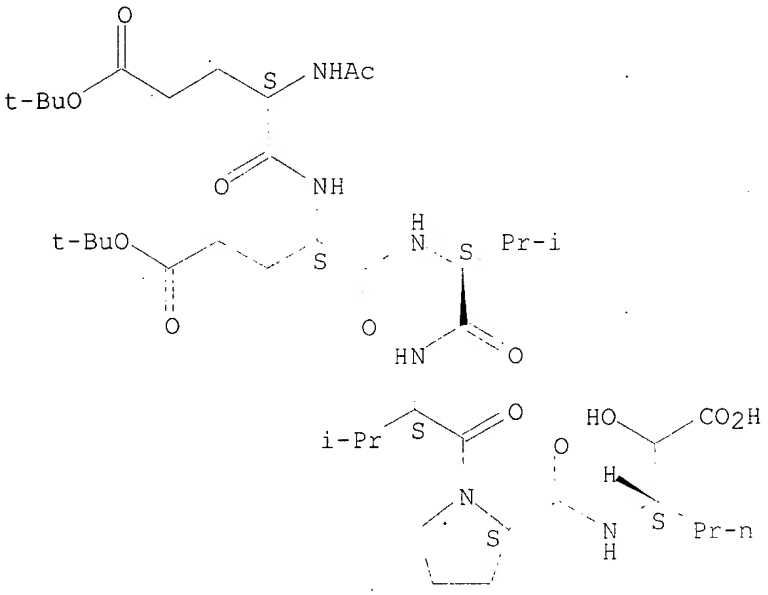
=> d sqide 17 tot

L7 ANSWER 1 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-43-4 REGISTRY
CN L-Prolinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-N-[(1S)-1-(carboxyhydroxymethyl)butyl]-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 5
NTE modified (modifications unspecified)

SEQ 1 EEVVP

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C41 H70 N6 O13
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

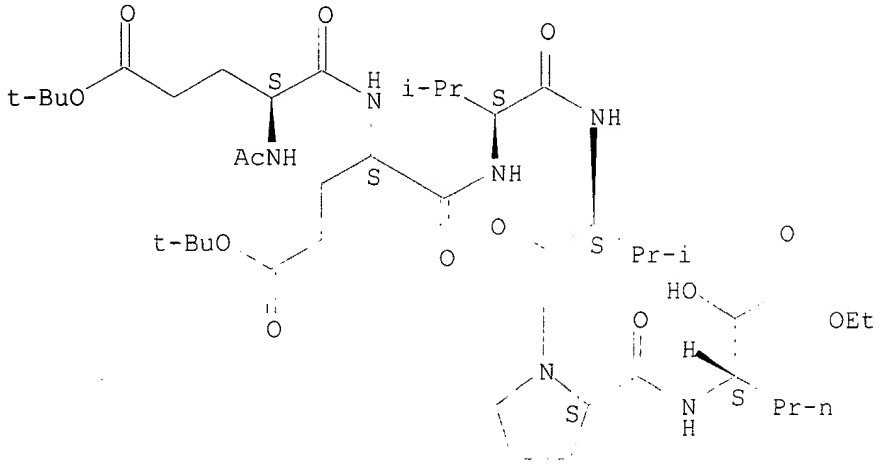
L7 ANSWER 2 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-42-3 REGISTRY
CN L-Prolinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-N-[(1S)-1-(2-ethoxy-1-hydroxy-2-oxoethyl)butyl]-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 5
NTE modified (modifications unspecified)

SEQ 1 EEVVP

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF C43 H74 N6 O13
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

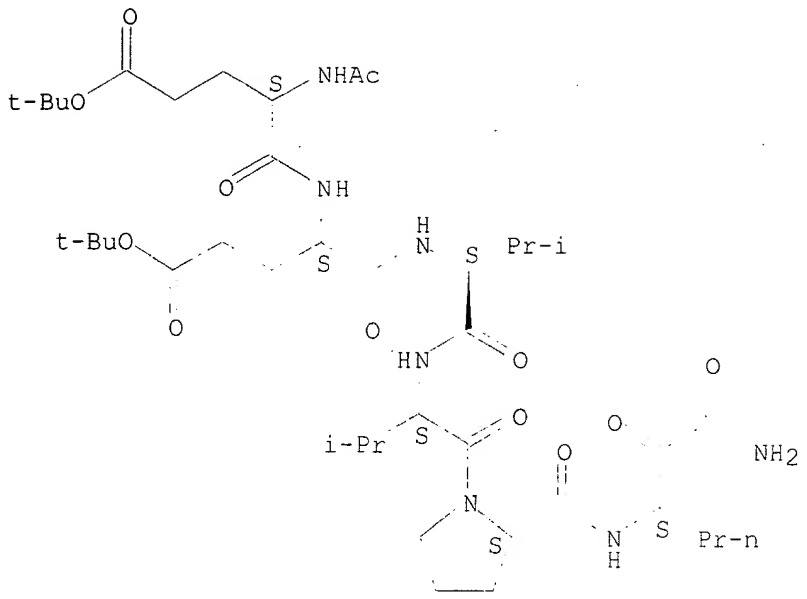
L7 ANSWER 3 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-40-1 REGISTRY
CN L-Prolinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-N-[(1S)-1-(aminooxoacetyl)butyl]-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 6
NTE modified

type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
uncommon	Nva-6	-	-
modification	Glu-1	-	1,1-dimethylethyl<t-Bu>
modification	Glu-2	-	1,1-dimethylethyl<t-Bu>

SEQ 1 EEVVPX

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C41 H69 N7 O12
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

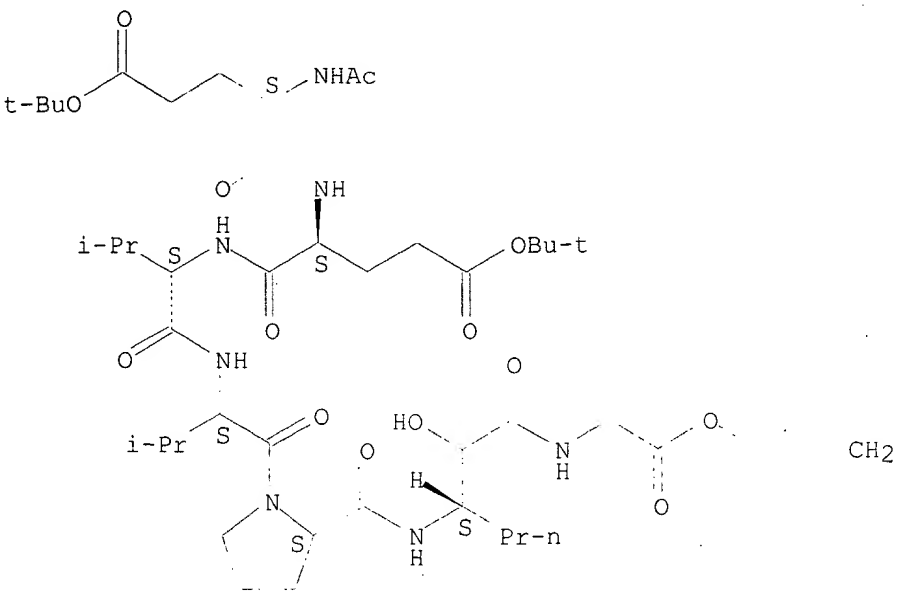
L7 ANSWER 4 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-37-6 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-hydroxyhexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propenyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------	description
uncommon	Oaa-6 - -	

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H77 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 5 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-31-0 REGISTRY
CN Glycinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2R,3S)-3-amino-2-hydroxyhexanoyl-N-2-propenyl-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

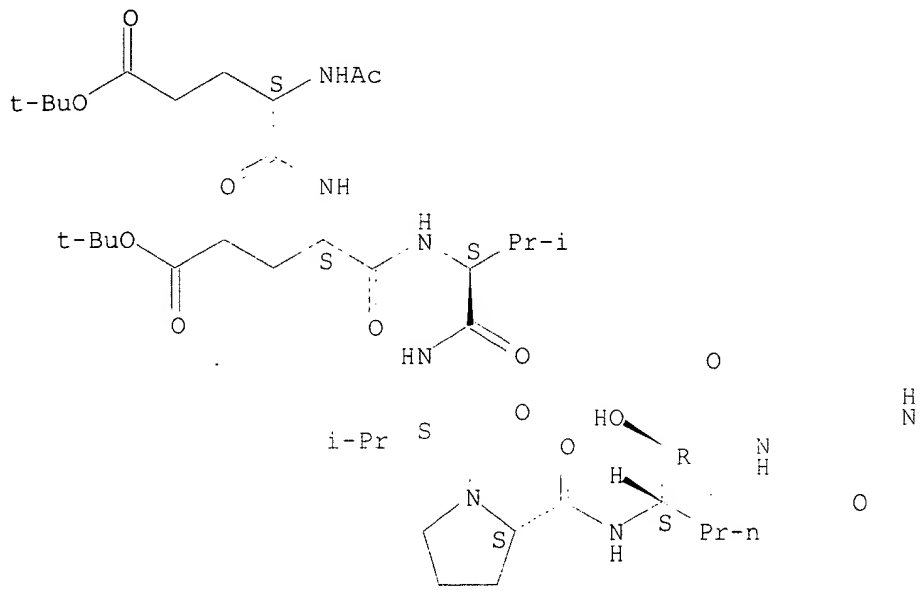
type	------ location ------	description
uncommon	Oaa-6	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H78 N8 O13
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 6 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-29-6 REGISTRY
CN Glycinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2S,3S)-3-amino-2-hydroxyhexanoyl-N-2-propenyl-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

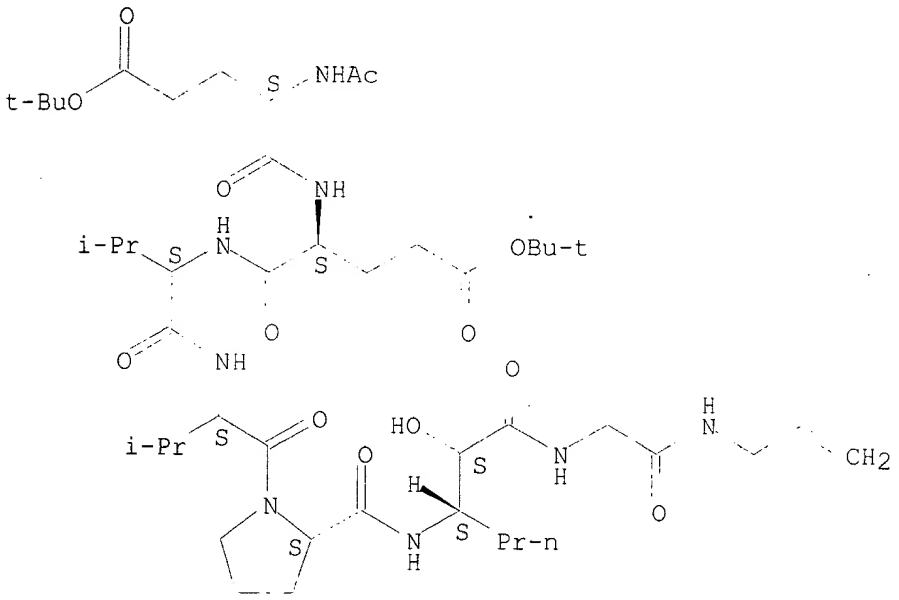
---

type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H78 N8 O13
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 7 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-27-4 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2R,3S)-3-amino-2-hydroxyhexanoyl-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

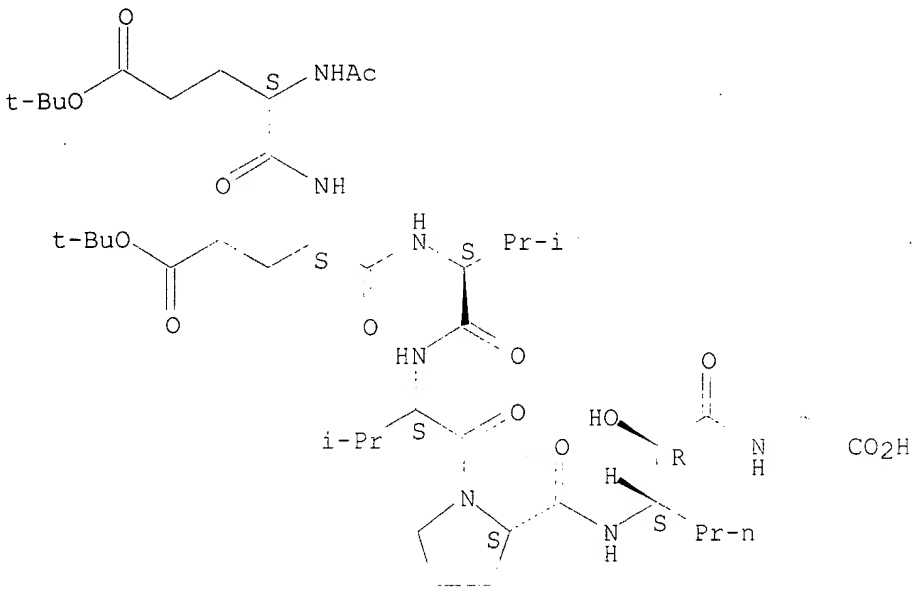
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C43 H73 N7 O14

SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

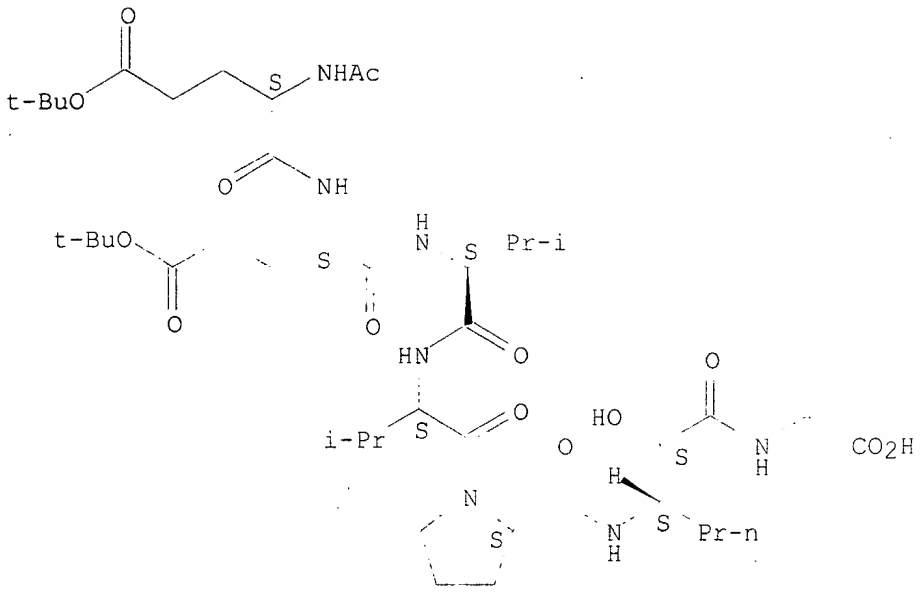
L7 ANSWER 8 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-25-2 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2S,3S)-3-amino-2-hydroxyhexanoyl-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------	description
uncommon	Oaa-6	- -

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C43 H73 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

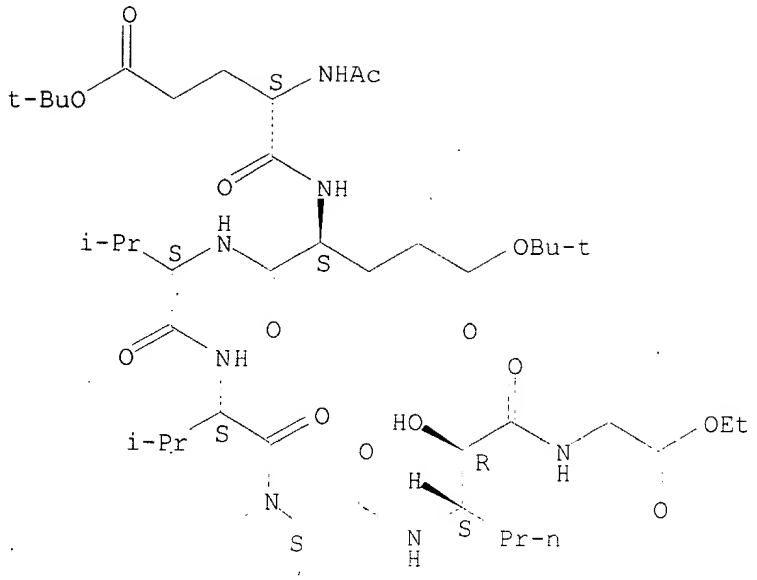
L7 ANSWER 9 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-23-0 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2R,3S)-3-amino-2-hydroxyhexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-ethyl ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C45 H77 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

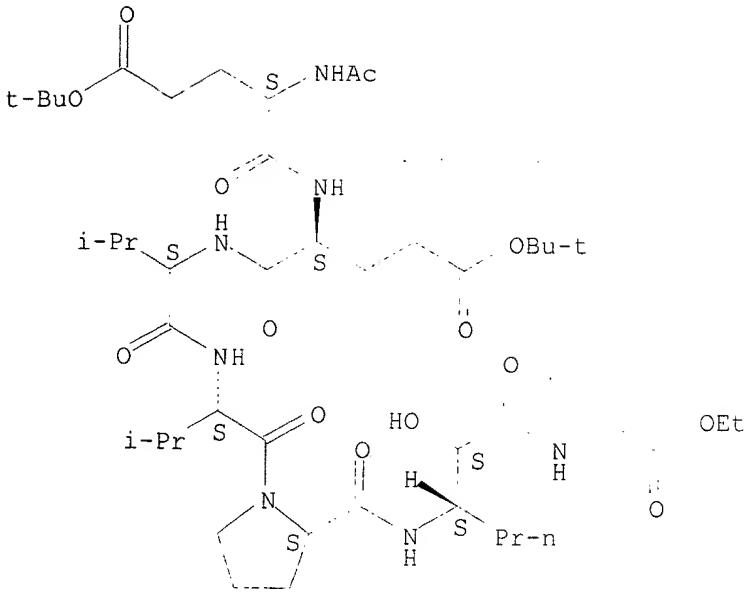
L7 ANSWER 10 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393525-21-8 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(2S,3S)-3-amino-2-hydroxyhexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-ethyl ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C45 H77 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

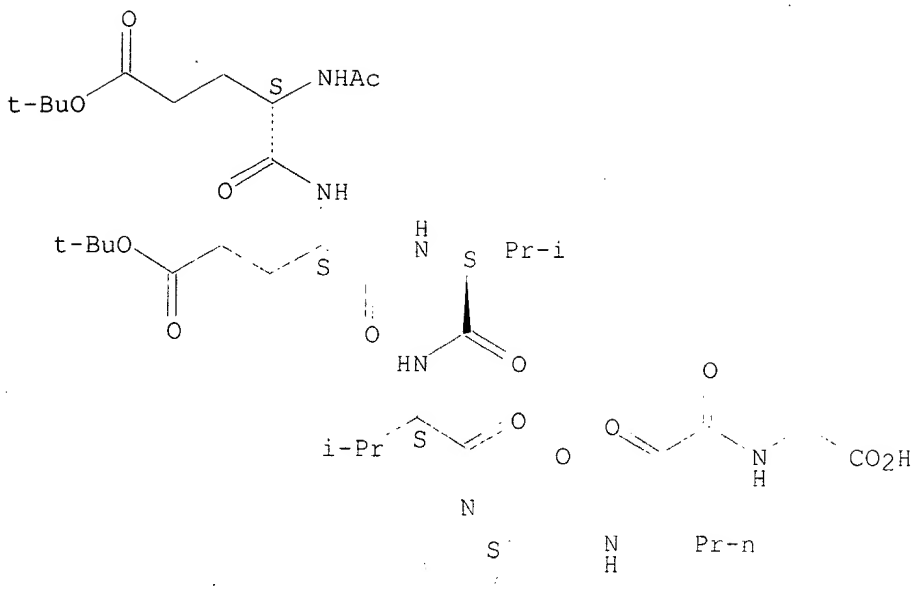
L7 ANSWER 11 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393521-13-6 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C43 H71 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

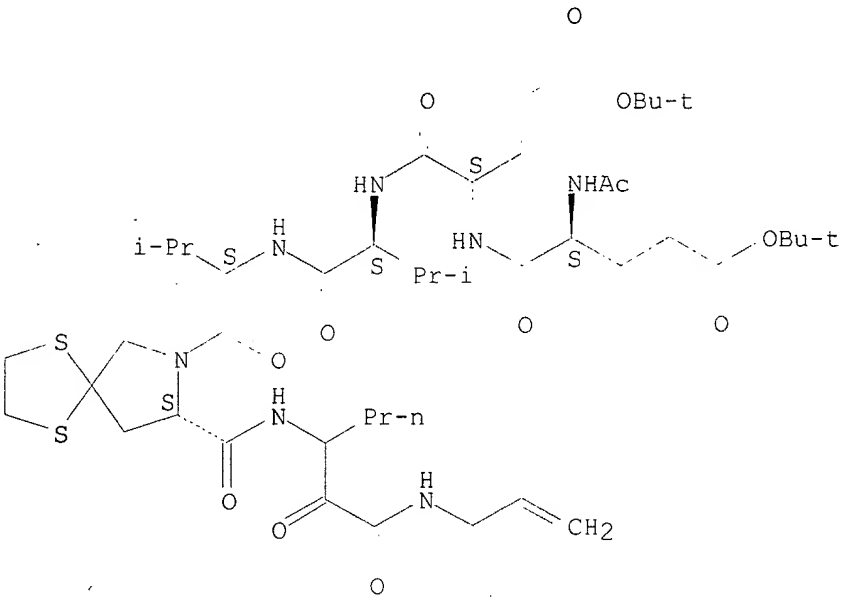
L7 ANSWER 12 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-87-1 REGISTRY
CN 1,4-Dithia-7-azaspiro[4.4]nonane-8-carboxamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-N-[1-[oxo(2-propenylamino)acetyl]butyl]-, bis(1,1-dimethylethyl) ester, (8S)- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 5
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Aaa-5	-	-

SEQ 1 EEVVX

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H75 N7 O12 S2
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 13 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-83-7 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-(4R)-4-[[(1,1-dimethylethyl)dimethylsilyl]oxy]-L-prolyl-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propenyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

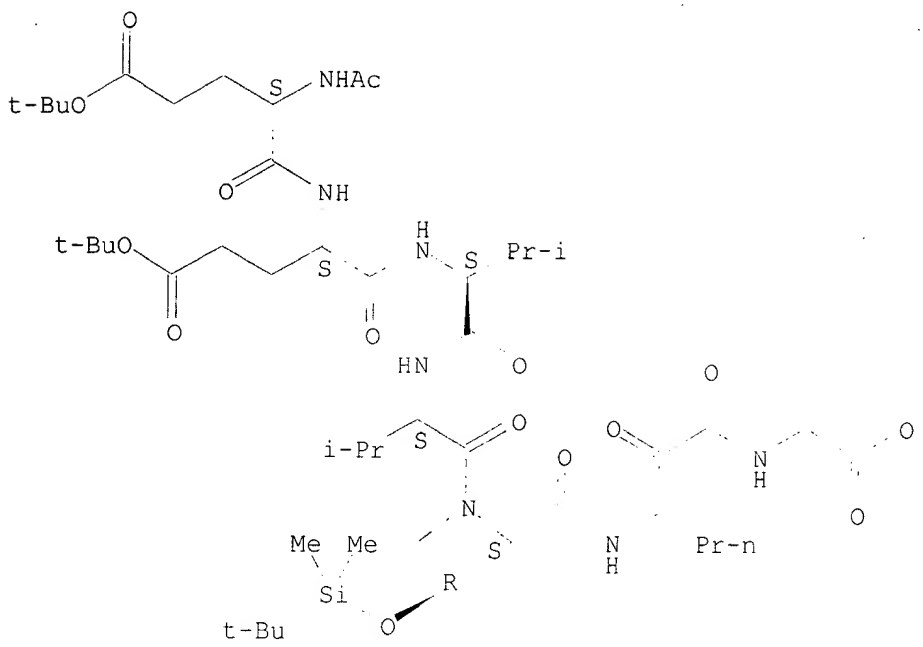
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C52 H89 N7 O15 Si
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

$=CH_2$

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 14 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-81-5 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-(4R)-4-(1,1-dimethylethoxy)-L-prolyl-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propenyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------	description

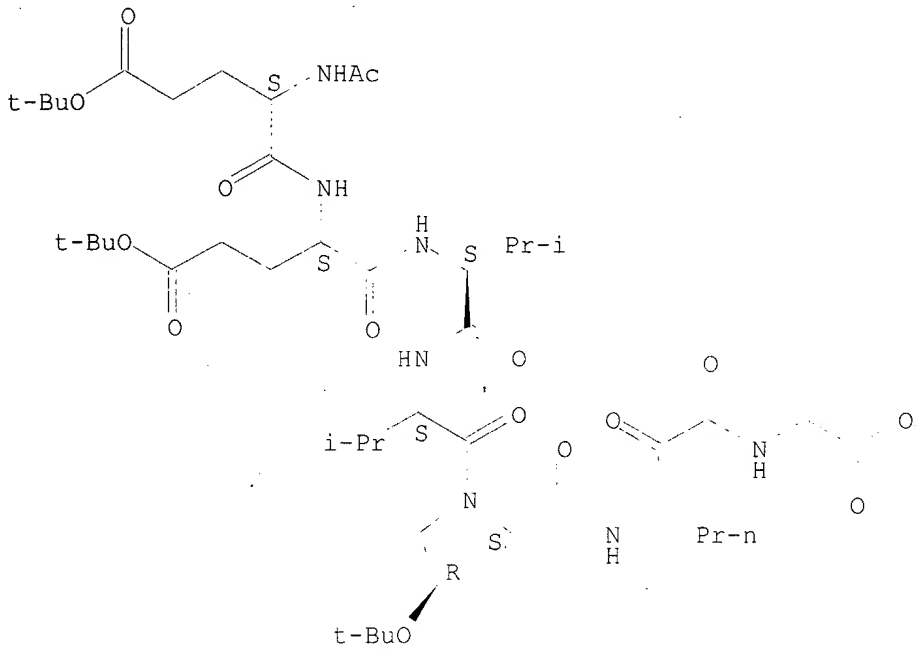
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C50 H83 N7 O15
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

$=CH_2$

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

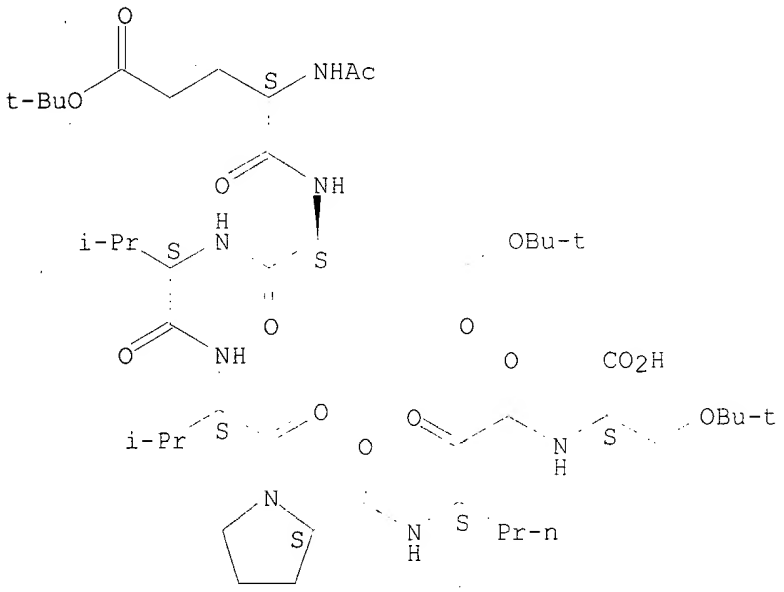
L7 ANSWER 15 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-63-3 REGISTRY
CN L-Serine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-O-(1,1-dimethylethyl)-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------	description
uncommon	Oaa-6	- -

SEQ 1 EEVVPXS

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C48 H81 N7 O15
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 16 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-61-1 REGISTRY
CN L-Serine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-O-(1,1-dimethylethyl)-,

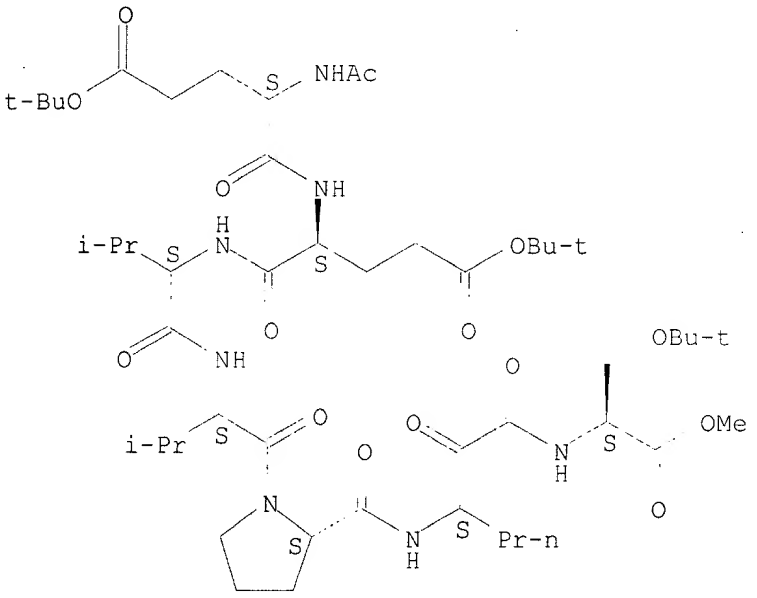
1,2-bis(1,1-dimethylethyl) 7-methyl ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXS

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C49 H83 N7 O15
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

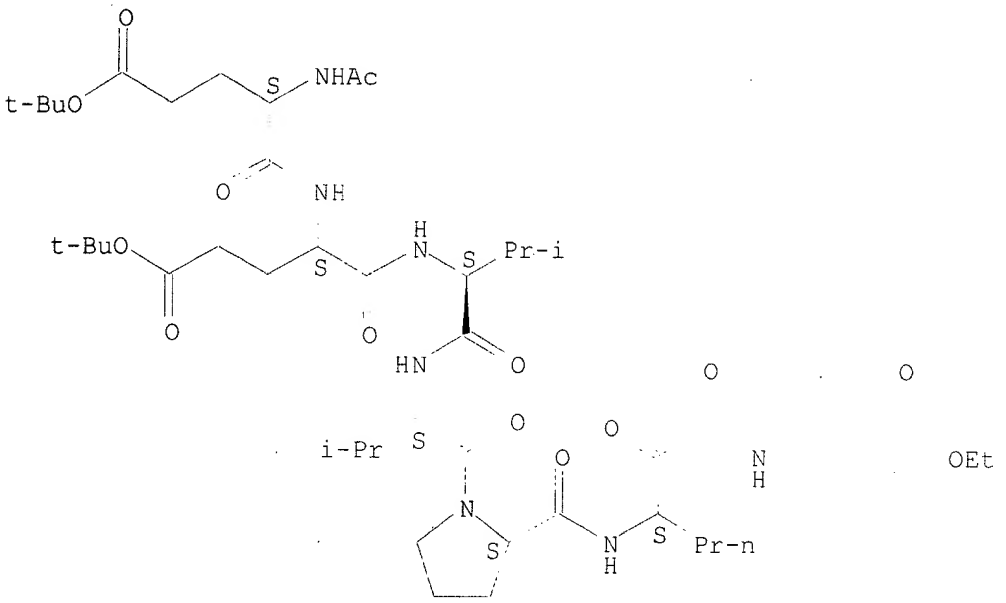
L7 ANSWER 17 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-51-9 REGISTRY
CN .beta.-Alanine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-ethyl ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

type	------ location ------		description
uncommon	Oaa-6	-	-
uncommon	Bal-7	-	-

SEQ 1 EEVVPXX

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H77 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 18 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-49-5 REGISTRY
CN L-Serinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-N-methyl-O-(phenylmethyl)-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

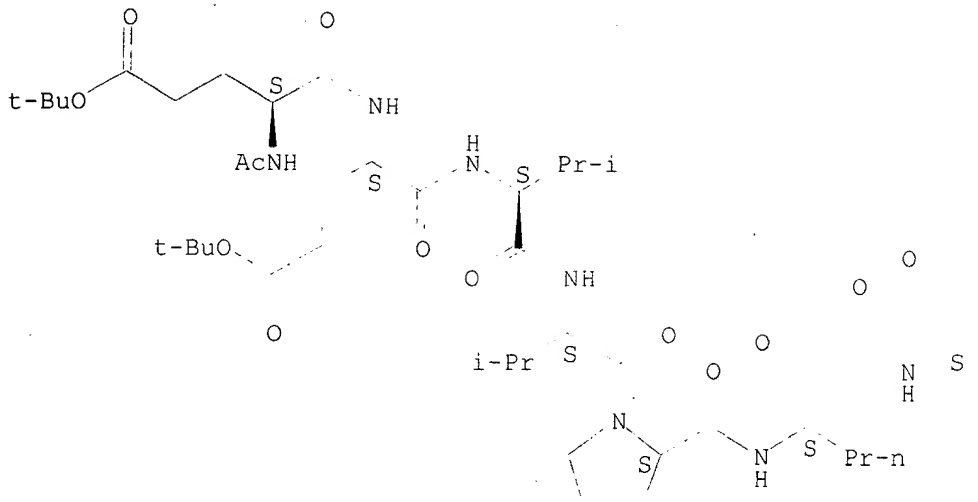
type	location	description
uncommon	Oaa-6 - -	

SEQ 1 EEVVPXS

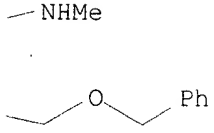
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C52 H82 N8 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 19 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-33-7 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-(4R)-4-[[(phenylmethoxy)carbonyl]amino]-L-prolyl-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propenyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

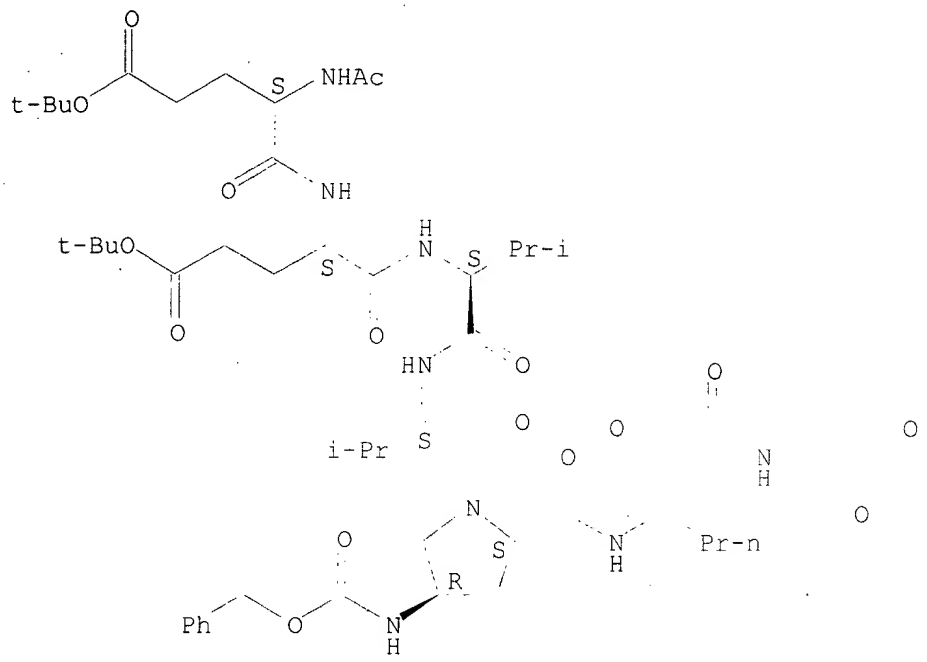
type	------ location ------	description
uncommon	Oaa-6	- -

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C54 H82 N8 O16
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

$=CH_2$

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 20 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-19-9 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propynyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

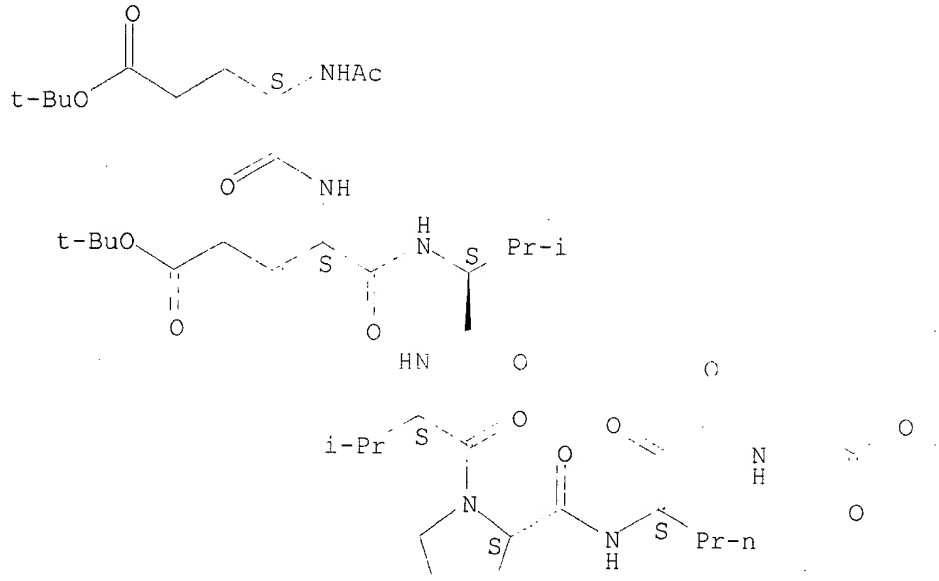
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H73 N7 O14.
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



—C≡CH

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 21 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393520-05-3 REGISTRY
CN Glycinamide, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-N-2-propenyl-, bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

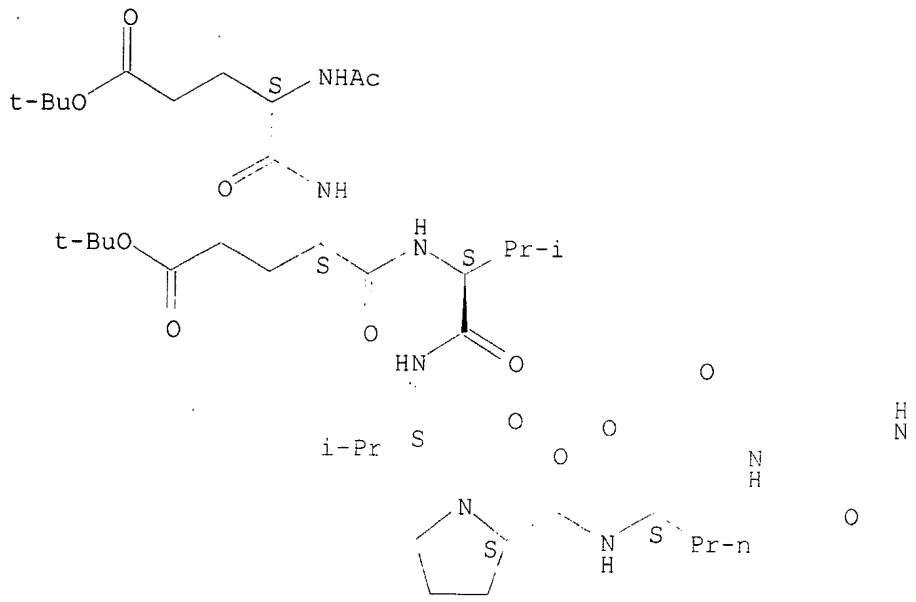
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C46 H76 N8 O13
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 22 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393519-93-2 REGISTRY
CN Glycine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-oxohexanoyl-, 1,2-bis(1,1-dimethylethyl) 7-(2-propenyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 7
NTE modified (modifications unspecified)

---

type	------ location ------	description
uncommon	Oaa-6 - -	

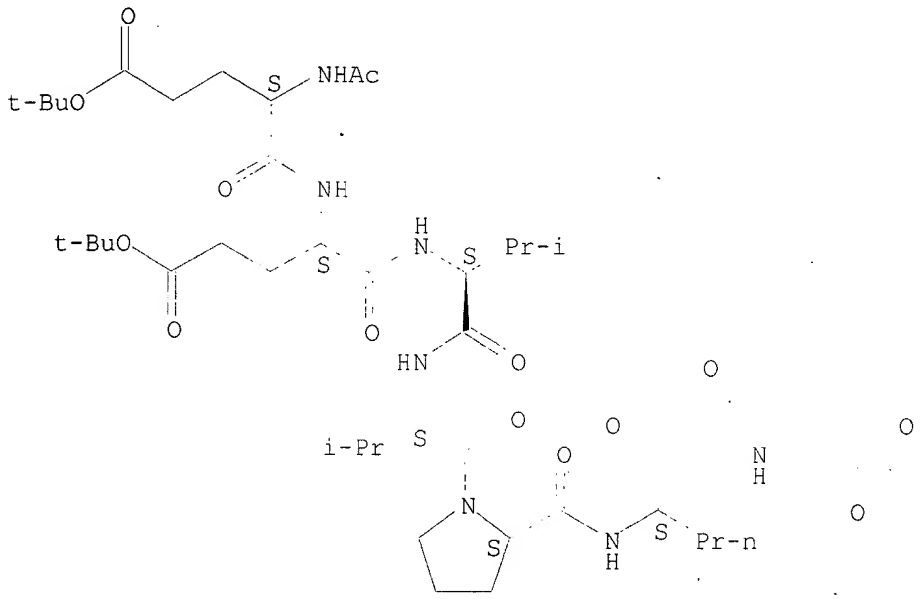
SEQ 1 EEVVPXG

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF C46 H75 N7 O14
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

$=CH_2$

PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 23 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393513-23-0 REGISTRY
CN L-Serine, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-L-prolyl-(3S)-3-amino-2-[[[(diphenylmethyl)amino]carbonyl]hydrazono]hexanoyl glycyl-L-methionyl-O-(1,1-dimethylethyl)-L-seryl-O-(1,1-dimethylethyl)-L-tyrosyl-O-(1,1-dimethylethyl)-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 11
NTE modified (modifications unspecified)

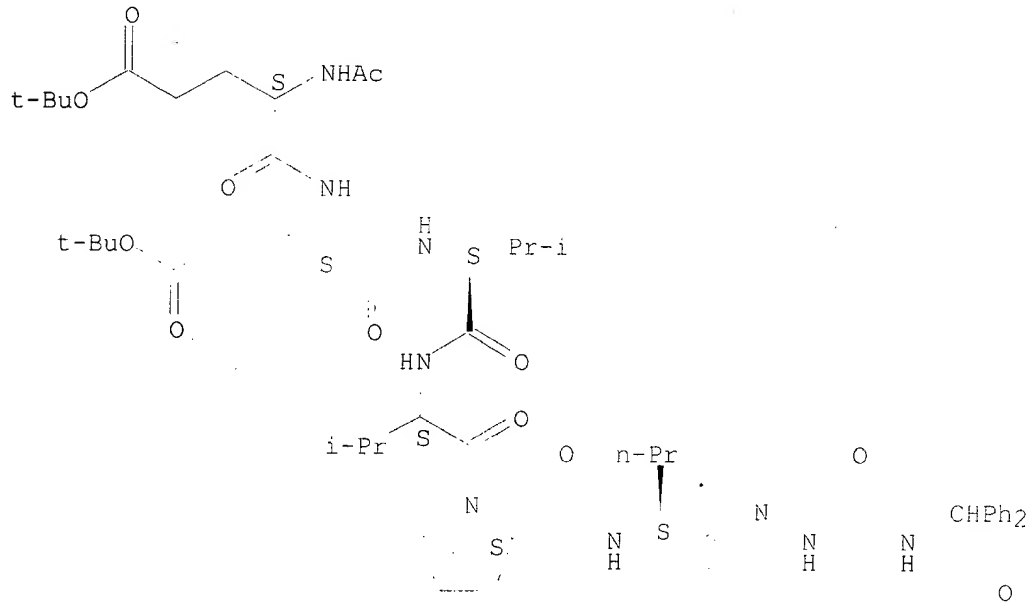
type	------ location ------		description
uncommon	Oaa-6	-	-

SEQ 1 EEVVPXGMSY S

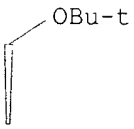
RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C89 H136 N14 O21 S
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.
Double bond geometry unknown.

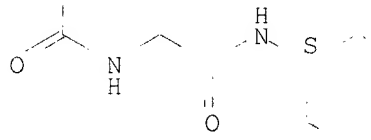
PAGE 1-A



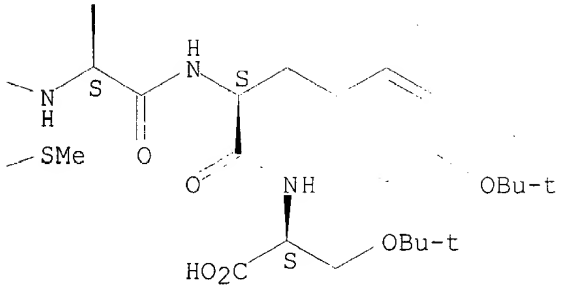
PAGE 1-B



PAGE 2-A



PAGE 2-B



1 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
1 REFERENCES IN FILE CAPLUS (1957 TO DATE)

L7 ANSWER 24 OF 24 REGISTRY COPYRIGHT 2003 ACS
RN 393513-18-3 REGISTRY
CN L-Proline, N-acetyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-valyl-L-valyl-, 1,2-bis(1,1-dimethylethyl) ester (9CI) (CA INDEX NAME)

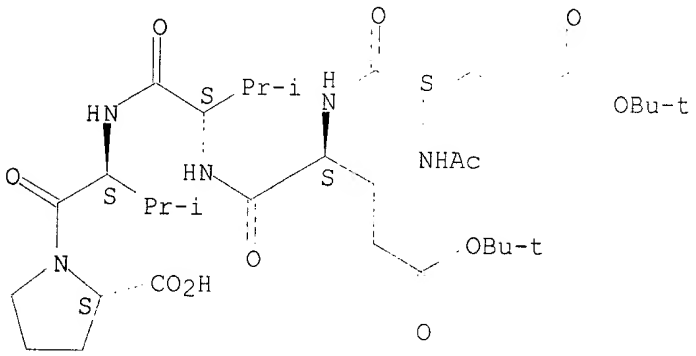
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 5
NTE modified

type	------ location ------		description
terminal mod.	Glu-1	-	N-acetyl
modification	Glu-1	-	1,1-dimethylethyl<t-Bu>
modification	Glu-2	-	1,1-dimethylethyl<t-Bu>

SEQ 1 EEVVP

RELATED SEQUENCES AVAILABLE WITH SEQLINK
MF C35 H59 N5 O11
SR CA
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

2 REFERENCES IN FILE CA (1957 TO DATE)
1 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA
2 REFERENCES IN FILE CAPLUS (1957 TO DATE)

0 qs then
0 search missing
middle 3 mers - too broad

```
=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 16:51:55 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)
```



```
FILE COVERS 1907 - 4 Jun 2003  VOL 138 ISS 23
FILE LAST UPDATED: 3 Jun 2003  (20030603/ED)

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.
```

Work discloses enough

```
=>
=>

=> d stat que l9
L4          740 SEA FILE=REGISTRY ABB=ON  PLU=ON  EEVVP/SQSP
L5       191606 SEA FILE=REGISTRY ABB=ON  PLU=ON  SYS/SQSP OR HYS/SQSP OR
                DYS/SQSP
L6          140 SEA FILE=REGISTRY ABB=ON  PLU=ON  L4 AND L5
L7           69 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L6
L8           68 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L7 NOT LIM?/AU,IN
L9           26 SEA FILE=HCAPLUS ABB=ON  PLU=ON  L8 NOT (2003 OR 2002 OR
                2001)/PY
```

5-mer

3-Mers (3 options)

8/11 mers

```
=>
=>

=> d ibib abs hitrn l9 1-26

L9   ANSWER 1 OF 26  HCAPLUS  COPYRIGHT 2003 ACS
ACCESSION NUMBER:         2001:108631  HCAPLUS
DOCUMENT NUMBER:          134:173731
TITLE:                    Sequence and analysis of chromosome 1 of the plant
                          Arabidopsis thaliana
AUTHOR(S):                Theologis, Athanasios; Ecker, Joseph R.; Palm, Curtis
                          J.; Federspiel, Nancy A.; Kaul, Samir; White, Owen;
                          Alonso, Jose; Altafi, Hootan; Araujo, Rina; Bowman,
                          Cheryl L.; Brooks, Shelise Y.; Buehler, Eugen; Chan,
                          April; Chao, Qimin; Chen, Huaming; Cheuk, Rosa F.;
                          Chin, Christina W.; Chung, Mike K.; Conn, Lane;
                          Conway, Aaron B.; Conway, Andrew R.; Creasy, Todd H.;
                          Dewar, Ken; Dunn, Patrick; Etgu, Pelin; Fedlblyum,
                          Tamara V.; Feng, Jidong; Fong, Betty; Fujii, Claire
                          Y.; Gill, John E.; Goldsmith, Andrew D.; Haas, Brian;
                          Hansen, Nancy F.; Hughes, Beth; Hulzar, Lucas; Hunter,
                          Johnathan L.; Jenkins, Jennifer; Johnson-Hopson,
                          Chanda; Khan, Shehnaz; Khaykin, Elizabeth; Kim,
                          Christopher J.; Koo, Hean L.; Kremenetskala, Irina;
                          Kurtz, David B.; Dwan, Andrea; Lam, Bao;
```

Langin-Hooper, Stephanie; Lee, Andrew; Lee, Jeong M.; Lenz, Catherine A.; Li, Joycelyn H.; Li, Yaping; Lin, Xiaoying; Liu, Shirley X.; Liu, Zhaoying A.; Luros, Jason S.; Malti, Rama; Marziall, Andre; Militscher, Jennifer; Miranda, Molly; Nguyen, Michelle; Nierman, William C.; Osborne, Brian I.; Pal, Grace; Peterson, Jeremy; Pham, Paul K.; Rizzo, Michael; Rooney, Timothy; Rowley, Don; Sakano, Hitomi; Salzberg, Steven L.; Schwartz, Jody R.; Shinn, Paul; Southwick, Audrey M.; Sun, Hui; Tallon, Luke J.; Tambunga, Gabrial; Toriumi, Mitsue J.; Town, Christopher D.; Utterback, Teresa; Van Aken, Susan; Vaysberg, Maria; Vysotskala, Valentina S.; Walker, Michelle; Wu, Dongying; Yu, Guixia; Fraser, Claire M.; Venter, J. Craig; Davis, Ronald W.

CORPORATE SOURCE: Plant Gene Expression Center/USDA-U.C. Berkeley, Albany, CA, 94710, USA

SOURCE: Nature (London) (2000), 408(6814), 816-820
CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The genome of the model plant Arabidopsis thaliana has been sequenced by an international collaboration, The Arabidopsis Genome Initiative. The complete sequence of the largest chromosome, chromosome 1, is reported in two contigs of around 14.2 and 14.6 megabases. The contigs extend from the telomeres to the centromeric borders, regions rich in transposons, retrotransposons and repetitive elements such as the 180-bp repeat. The chromosome represents 25% of the genome and contains about 6850 open reading frame,s, 236 tRNAs, and 12 small nuclear RNAs. There are two clusters of tRNA genes at different places on the chromosome. One consts of 27 tRNAPro genes and the other contains 27 tandem repeats of tRNATyr-tRNATyr-tRNASer genes. Chromosome 1 contains about 300 gene families with clustered duplications. There are also many repeat elements, representing 8% of the sequence.

IT **324098-98-8**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; sequence and anal. of chromosome 1 of the plant Arabidopsis thaliana)

L9 ANSWER 2 OF 26 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2000:690495 HCAPLUS
DOCUMENT NUMBER: 133:347455
TITLE: Drosophila D-Titin is required for myoblast fusion and skeletal muscle striation

AUTHOR(S): Zhang, Yong; Featherstone, David; Davis, Warren; Rushton, Emma; Broadie, Kendal

CORPORATE SOURCE: Department of Biology, University of Utah, Salt Lake City, UT, 84112-0840, USA

SOURCE: Journal of Cell Science (2000), 113(17), 3103-3115
CODEN: JNCSAI; ISSN: 0021-9533

PUBLISHER: Company of Biologists Ltd.
DOCUMENT TYPE: Journal
LANGUAGE: English

AB An ethylmethane sulfonate (EMS) mutagenesis of D. melanogaster aimed at discovering novel genes essential for neuromuscular development identified 6 embryonic lethal alleles of 1 genetic locus on the 3rd chromosome at 62C. Two addnl. lethal P element insertion lines, l(3)S02001 and l(3)j1D7, failed to complement each other and each of the 6 EMS alleles. Anal. of genomic sequence bracketing the 2 insertion sites predicted a protein of 16,215 amino acid residues, encoded by a 70 kb genomic region. This sequence includes the recently characterized kettin, and includes all

known partial D-Titin sequences. We call the genetic locus, which encodes both D-Titin and kettin, D-Titin. D-Titin has 53 repeats of the Ig C2 domain, 6 repeats of the fibronectin type III domain and two large PEVK domains. Kettin appears to be the N-terminal 33% of D-Titin, presumably expressed via alternative splicing. Phenotype assays on the allelic series of D-Titin mutants demonstrated that D-Titin has an unsuspected function in myoblast fusion during myogenesis and, 2nd, D-Titin later serves to organize myofilaments into the highly ordered arrays underlying skeletal muscle striation. We propose that D-Titin is instrumental in the development of the 2 defining features of striated muscle: the formation of multi-nucleate syncytia and the organization of actin-myosin filaments into striated arrays.

IT **306331-52-2**, Titin (Drosophila melanogaster)
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; Drosophila titin sequence and role in myoblast fusion and skeletal muscle striation)

REFERENCE COUNT: 48 THERE ARE 48 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 3 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:613167 HCAPLUS
DOCUMENT NUMBER: 133:218310
TITLE: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae
AUTHOR(S): Heidelberg, John F.; Elsen, Jonathan A.; Nelson, William C.; Clayton, Rebecca A.; Gwinn, Michelle L.; Dodson, Robert J.; Haft, Daniel H.; Hickey, Erin K.; Peterson, Jeremy D.; Umayam, Lowell; Gill, Steven R.; Nelson, Karen E.; Read, Timothy D.; Tettelin, Herve; Richardson, Delwood; Ermolaeva, Maria D.; Vamathevan, Jessica; Bass, Steven; Qin, Haiying; Dragoi, Ioana; Sellers, Patrick; McDonald, Lisa; Utterback, Teresa; Fleishmann, Robert D.; Nierman, William C.; White, Owen; Salzberg, Steven L.; Smith, Hamilton O.; Colwell, Rita R.; Mekalanos, John J.; Venter, J. Craig; Fraser, Claire M.
CORPORATE SOURCE: The Institute for Genomic Research, Rockville, MD, 20850, USA
SOURCE: Nature (London) (2000), 406(6795), 477-483
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Nature Publishing Group
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The complete genomic sequence of the gram-neg., .gamma.-Proteobacterium Vibrio cholerae El Tor N16961 was detd. to be 4,033,460 bp. The genome consists of two circular chromosomes of 2,961,146 bp and 1,072,314 bp that together encode 3885 open reading frames. The vast majority of recognizable genes for essential cell functions (such as DNA replication, transcription, translation, and cell-wall biosynthesis) and pathogenicity (for example, toxins, surface antigens, and adhesins) are located on the large chromosome. In contrast, the small chromosome contains a larger fraction (59%) of hypothetical genes compared with the large chromosome (42%), and also contains many more genes that appear to have origins other than the .gamma.-Proteobacteria. The small chromosome also carries a gene capture system (the integron island) and host 'addiction' genes that are typically found on plasmids; thus, the small chromosome may have originally been a megaplasmid that was captured by an ancestral Vibrio species. The V. cholerae genomic sequence provides a starting point for understanding how a free-living, environmental organism emerged to become a significant human bacterial pathogen.

IT **290391-05-8**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL

(Biological study)
(amino acid sequence; DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae)
REFERENCE COUNT: 50 THERE ARE 50 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 4 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:423003 HCAPLUS
DOCUMENT NUMBER: 133:291704
TITLE: Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity
AUTHOR(S): Freiburg, Alexandra; Trombitas, Karoly; Hell, Wolfgang; Cazorla, Olivier; Fougerousse, Francoise; Centner, Thomas; Kolmerer, Bernhard; Witt, Christian; Beckmann, Jaques S.; Gregorio, Carol C.; Granzier, Henk; Labeit, Siegfried
CORPORATE SOURCE: European Molecular Biology Laboratory, Institut fur Anasthesiologie und Operative Intensivmedizin, Heidelberg, 60012, Germany
SOURCE: Circulation Research (2000), 86(11), 1114-1121
CODEN: CIRUAL; ISSN: 0009-7330
PUBLISHER: Lippincott Williams & Wilkins
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Titins are megadalton-sized filamentous polypeptides of vertebrate striated muscle. The I-band region of titin underlies the myofibrillar passive tension response to stretch. Here, we show how titins with highly diverse I-band structures and elastic properties are expressed from a single gene. The differentially expressed tandem-Ig, PEVK, and N2B spring elements of titin are coded by 158 exons, which are contained within a 106-kb genomic segment and are all subject to tissue-specific skipping events. In ventricular heart muscle, exons 101 kb apart are joined, leading to the exclusion of 155 exons and the expression of a 2.97-MDa cardiac titin N2B isoform. The atria of mammalian hearts also express larger titins by the exclusion of 90 to 100 exons (cardiac N2BA titin with 3.3 MDa). In the soleus and psoas skeletal muscles, different exon-skipping pathways produce titin transcripts that code for 3.7- and 3.35-MDa titin isoforms, resp. Mech. and structural studies indicate that the exon-skipping pathways modulate the fractional extensions of the tandem Ig and PEVK segments, thereby influencing myofibrillar elasticity. Within the mammalian heart, expression of different levels of N2B and N2BA titins likely contributes to the elastic diversity of atrial and ventricular myofibrils.
IT **300595-83-9**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; partial sequence of human titin)
REFERENCE COUNT: 17 THERE ARE 17 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 5 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:230405 HCAPLUS
DOCUMENT NUMBER: 132:304167
TITLE: The genome sequence of Drosophila melanogaster
AUTHOR(S): Adams, Mark D.; Celniker, Susan E.; Holt, Robert A.; Evans, Cheryl A.; Gocayne, Jeannine D.; Amanatides, Peter G.; Scherer, Steven E.; Li, Peter W.; Hoskins, Roger A.; Galle, Richard F.; George, Reed A.; Lewis, Suzanna E.; Richards, Stephen; Ashburner, Michael; Henderson, Scott N.; Sutton, Granger G.; Wortman, Jennifer R.; Yandell, Mark D.; Zhang, Qing; Chen, Lin X.; Brandon, Rhonda C.; Rogers, Yu-Hui C.; Blazej,

Robert G.; Champe, Mark; Pfeiffer, Barret D.; Wan, Kenneth H.; Doyle, Clare; Baxter, Evan G.; Helt, Gregg; Nelson, Catherine R.; Miklos, George L. Gabor; Abril, Josep F.; Agbayani, Anna; An, Hui-Jin; Andrews-Pfannkoch, Cynthia; Baldwin, Danita; Ballew, Richard M.; Basu, Anand; Baxendale, James; Bayraktaroglu, Leyla; Beasley, Ellen M.; Beeson, Karen Y.; Benos, P. V.; Berman, Benjamin P.; Bhandari, Deepali; Bolshakov, Slava; Borkova, Dana; Botchan, Michael R.; Bouck, John; Brokstein, Peter; Brottier, Phillipe; Burtis, Kenneth C.; Busam, Dana A.; Butler, Heather; Cadieu, Edouard; Center, Angela; Chandra, Ishwar; Cherry, J. Michael; Cawley, Simon; Dahlke, Carl; Davenport, Lionel B.; Davies, Peter; De Pablos, Beatriz; Delcher, Arthur; Deng, Zuoming; Mays, Anne Deslattes; Dew, Ian; Dietz, Suzanne M.; Dodson, Kristina; Doup, Lisa E.; Downes, Michael; Dugan-Rocha, Shannon; Dunkov, Boris C.; Dunn, Patrick; Durbin, Kenneth J.; Evangelista, Carlos C.; Ferraz, Concepcion; Ferriera, Steven; Fleischmann, Wolfgang; Foster, Carl; Gabrielian, Andrei E.; Garg, Neha S.; Gelbart, William M.; Glasser, Ken; Glodek, Anna; Gong, Fangcheng; Gorrell, J. Harley; Gu, Zhiping; Guan, Ping; Harris, Michael; Harris, Nomi L.; Harvey, Damon; Heiman, Thomas J.; Hernandez, Judith R.; Houck, Jarrett; Hostin, Damon; Houston, Kathryn A.; Howland, Timothy J.; Wei, Ming-Hui; Ibegwam, Chinyere; Jalali, Mena; Kalush, Francis; Karpen, Gary H.; Ke, Zhaoxi; Kennison, James A.; Ketchum, Karen A.; Kimmel, Bruce E.; Kodira, Chinnappa D.; Kraft, Cheryl; Kravitz, Saul; Kulp, David; Lai, Zhongwu; Lasko, Paul; Lei, Yiding; Levitsky, Alexander A.; Li, Jiayin; Li, Zhenya; Liang, Yong; Lin, Xiaoying; Liu, Xiangjun; Mattei, Bettina; McIntosh, Tina C.; McLeod, Michael P.; McPherson, Duncan; Merkulov, Gennady; Milshina, Natalia V.; Mobarry, Clark; Morris, Joe; Moshrefi, Ali; Mount, Stephen M.; Moy, Mee; Murphy, Brian; Murphy, Lee; Muzny, Donna M.; Nelson, David L.; Nelson, David R.; Nelson, Keith A.; Nixon, Katherine; Nusskern, Deborah R.; Pacleb, Joanne M.; Palazzolo, Michael; Pittman, Gjange S.; Pan, Sue; Pollard, John; Puri, Vinita; Reese, Martin G.; Reinert, Knut; Remington, Karin; Saunders, Robert D. C.; Scheeler, Frederick; Shen, Hua; Shue, Bixiang Christopher; Siden-Kiamos, Inga; Simpson, Michael; Skupski, Marian P.; Smith, Tom; Spier, Eugene; Spradling, Allan C.; Stapleton, Mark; Strong, Renee; Sun, Eric; Svirskas, Robert; Tector, Cyndee; Turner, Russell; Venter, Eli; Wang, Aihui H.; Wang, Xin; Wang, Zhen-Yuan; Wassarman, David A.; Weinstock, George M.; Weissenbach, Jean; Williams, Sherita M.; Woodage, Trevor; Worley, Kim C.; Wu, David; Yang, Song; Yao, Q. Alison; Ye, Jane; Yeh, Ru-Fang; Zaveri, Jayshree S.; Zhan, Ming; Zhang, Guangren; Zhao, Qi; Zheng, Liansheng; Zheng, Xiangqun H.; Zhong, Fei N.; Zhong, Wenyan; Zhou, Xiaojun; Zhu, Shiaoping; Zhu, Xiaohong; Smith, Hamilton O.; Gibbs, Richard A.; Myers, Eugene W.; Rubin, Gerald M.; Venter, J. Craig

CORPORATE SOURCE: Celera Genomics, Rockville, MD, 20850, USA

SOURCE: Science (Washington, D. C.) (2000), 287(5461), 2185-2195
CODEN: SCIEAS; ISSN: 0036-8075

PUBLISHER: American Association for the Advancement of Science

DOCUMENT TYPE: Journal
LANGUAGE: English
AB The fly Drosophila melanogaster is one of the most intensively studied organisms in biol. and serves as a model system for the investigation of many developmental and cellular processes common to higher eukaryotes, including humans. The nucleotide sequence was detd. of nearly all of the .apprx.120-megabase euchromatic portion of the Drosophila genome using a whole-genome shotgun sequencing strategy supported by extensive clone-based sequence and a high-quality bacterial artificial chromosome phys. map. Efforts are under way to close the remaining gaps; however, the sequence is of sufficient accuracy and contiguity to be declared substantially complete and to support an initial anal. of genome structure and preliminary gene annotation and interpretation. The genome encodes .apprx.13,600 genes, somewhat fewer than the smaller Caenorhabditis elegans genome, but with comparable functional diversity. Access to supporting information on each gene is available through FlyBase at http://flybase.bio.indiana.edu and through Celera at www.celera.com; the sequences are deposited in GenBank with Accession Nos. AE002566-AE003403. [This abstr. record is one of 4 records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].
IT **262986-25-4**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; genome sequence of Drosophila melanogaster)
REFERENCE COUNT: 89 THERE ARE 89 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 6 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:114404 HCAPLUS
DOCUMENT NUMBER: 132:176625
TITLE: Multidrug resistance protein MRP and cDNA and their use in therapy and drug screening
INVENTOR(S): Deeley, Roger G.; Cole, Susan P. C.
PATENT ASSIGNEE(S): Queen's University At Kingston, Can.
SOURCE: U.S., 78 pp., Cont.-in-part of U.S. Ser. No. 407,207.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 8
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6025473	A	20000215	US 1995-461384	19950605
US 5489519	A	19960206	US 1993-141893	19931026
US 6063621	A	20000516	US 1995-407207	19950320

PRIORITY APPLN. INFO.: US 1992-966923 B2 19921027
US 1993-29340 B2 19930308
US 1993-141893 A2 19931026
US 1995-407207 A2 19950320

AB The present invention relates to isolation of human multidrug resistance protein MRP and cDNA and their use in therapy and drug screening. The cDNA for MRP which is capable of conferring multidrug resistance on cells not expressing P-glycoprotein was cloned from H69AR cells. Northern blot anal. indicated that MRP was produced in relatively high levels in lung, testis, and peripheral blood mononuclear cells. The protein sequence anal. showed MRP belongs to ATP-binding cassette superfamily of membrane transport proteins but the human MRP gene was mapped to 16p13.1, indicating that it is not linked to either CFTR or MDR genes. Polyclonal and monoclonal antibodies were prepd. to MRP and tested in further MRP characterization and analytic assay. Expression of the MRP cDNA in drug-sensitive mammalian cells conferred multidrug resistance upon these

cells. The cDNAs for a variant form of human MRP (with changes of T2249 to C and G4040 to C in DNA and corresponding Leu685 to Ser and Arg1282 to Ala in protein) as well as for the mouse homolog of MRP were also cloned and sequenced. Diagnostic and treatment methods using the novel proteins, nucleic acids, antibodies and cell lines of the invention are also encompassed by the invention.

IT **179046-57-2**, Protein MRP (mouse testis multidrug resistance-associated)
RL: BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PROC (Process); USES (Uses)
(amino acid sequence; multidrug resistance protein MRP and cDNA and their use in therapy and drug screening)

REFERENCE COUNT: 66 THERE ARE 66 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 7 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:794245 HCAPLUS
DOCUMENT NUMBER: 132:19669
TITLE: Multidrug resistance protein MRP and cDNA and method for identifying cytotoxic agents for multidrug resistant cells
INVENTOR(S): Deeley, Roger G.; Cole, Susan Pc
PATENT ASSIGNEE(S): Queen's University at Kingston, Can.
SOURCE: U.S., 77 pp., Cont.-in-part of U. S. Ser. No. 407,207.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 8
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6001563	A	19991214	US 1995-463179	19950605
US 5489519	A	19960206	US 1993-141893	19931026
US 6063621	A	20000516	US 1995-407207	19950320

PRIORITY APPLN. INFO.: US 1992-966923 B2 19921027
US 1993-29340 B2 19930308
US 1993-141893 A2 19931026
US 1995-407207 A2 19950320

AB Novel protein MRP (multidrug resistance-assocd. protein) which is capable of conferring multidrug resistance on cells not expressing P-glycoprotein and nucleic acids encoding the novel protein are disclosed. Transformant cell lines which express the nucleic acid encoding MRP and their use in identification of agents cytotoxic to multidrug resistant cells are claimed. The cDNA for MRP was cloned from H69AR cells. Northern blot anal. indicated that MRP was produced in relatively high levels in lung, testis, and peripheral blood mononuclear cells. The human MRP gene was mapped to 16p13.1, indicating that it is not linked to either CFTR or MDR genes.a. Antibodies were prepd. to the protein. The protein was found to be glycosylated and to bind ATP. Expression of the MRP cDNA in drug-sensitive mammalian cells conferred multidrug resistance upon these cells. The cDNAs for a variant form of human MRP as well as for the mouse homolog of MRP were also cloned and sequenced.

IT **179046-57-2**, Protein MRP (mouse testis multidrug resistance-associated)
RL: ARG (Analytical reagent use); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); ANST (Analytical study); BIOL (Biological study); PROC (Process); USES (Uses)
(amino acid sequence; Multidrug resistance protein MRP and cDNA and method for identifying cytotoxic agents for multidrug resistant cells)

REFERENCE COUNT: 22 THERE ARE 22 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 8 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:722042 HCAPLUS
DOCUMENT NUMBER: 132:74206
TITLE: Characterization of cDNA clones selected by the GeneMark analysis from size-fractionated cDNA libraries from human brain
AUTHOR(S): Hirosawa, Makoto; Nagase, Takahiro; Ishikawa, Ken-Ichi; Kikuno, Reiko; Nomura, Nobuo; Ohara, Osamu
CORPORATE SOURCE: Kazusa DNA Research Institute, Chiba, 292-0812, Japan
SOURCE: DNA Research (1999), 6(5), 329-336
CODEN: DARSE8; ISSN: 1340-2838
PUBLISHER: Universal Academy Press
DOCUMENT TYPE: Journal
LANGUAGE: English
AB We have conducted a sequencing project of human cDNAs which encode large proteins in brain. For selection of cDNA clones to be sequenced in this project, cDNA clones have been exptl. examd. by in vitro transcription/translation prior to sequencing. In this study, we tested an alternative approach for picking up cDNA clones having a high probability of carrying protein coding region. This approach exploited 5'-end single-pass sequence data and the GeneMark program for assessing protein-coding potential, and allowed us to select 74 clones out of 14,804 redundant cDNA clones. The complete sequence data of these 74 clones revealed that 45% of them encoded proteins consisting of more than 500 amino acid residues while all the clones thus selected carried possible protein coding sequences as expected. The results indicated that the GeneMark anal. of 5'-end sequences of cDNAs offered us a simple and effective means to select cDNA clones with protein-coding potential although the sizes of the encoded proteins could not be predicted.
IT **253424-06-5**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; characterization of cDNA clones selected by GeneMark anal. from size-fractionated cDNA libraries from human brain)
REFERENCE COUNT: 11 THERE ARE 11 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 9 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:476579 HCAPLUS
DOCUMENT NUMBER: 131:238581
TITLE: Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
AUTHOR(S): Kikuno, Reiko; Nagase, Takahiro; Ishikawa, Ken-Ichi; Hirosawa, Makoto; Miyajima, Nobuyuki; Tanaka, Ayako; Kotani, Hirokazu; Nomura, Nobuo; Ohara, Osamu
CORPORATE SOURCE: Kazusa DNA Research Institute, Chiba, 292-0812, Japan
SOURCE: DNA Research (1999), 6(3), 197-205
CODEN: DARSE8; ISSN: 1340-2838
PUBLISHER: Universal Academy Press
DOCUMENT TYPE: Journal
LANGUAGE: English
AB To extend our cDNA project for accumulating basic information on unidentified human genes, we newly detd. the sequences of 100 cDNA clones from a set of size-fractionated human adult and fetal brain cDNA libraries, and predicted the coding sequences of the corresponding genes, named KIAA1019 to KIAA1118. The sequencing of these clones revealed that the av. size of the inserts and corresponding open reading frames were 5.0 kb and 2.6 kb (880 amino acid residues), resp. Database search of the predicted amino acid sequences classified 58 predicted gene products into the five functional categories, such as cell signaling/communication, cell

structure/motility, nucleic acid management, protein management and cell division. It was also found that, for 34 gene products, homologues were detected in the databases, which were similar in sequence through almost the entire regions. The chromosomal locations of the genes were detd. by using human-rodent hybrid panels unless their mapping data were already available in the public databases. The expression profiles of all the genes among 10 human tissues, 8 brain regions (amygdala, corpus callosum, cerebellum, caudate nucleus, hippocampus, substania nigra, subthalamic nucleus, and thalamus), spinal cord, fetal brain and fetal liver were also examd. by reverse transcription-coupled polymerase chain reaction, products of which were quantified by ELISA.

IT **244204-35-1**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; sequences of 100 new cDNA clones from human brain which code for large proteins in vitro)

REFERENCE COUNT: 16 THERE ARE 16 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 10 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:231174 HCAPLUS
DOCUMENT NUMBER: 130:247846
TITLE: Expression of multidrug resistance-associated protein MRP nucleic acid in cells to confer drug resistance
INVENTOR(S): Deeley, Roger G.; Cole, Susan P. C.
PATENT ASSIGNEE(S): Queen's University at Kingston, Can.
SOURCE: U.S., 82 pp., Cont.-in-part of U.S. Ser. No. 407,207.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 8
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5891724	A	19990406	US 1995-460907	19950605
US 5489519	A	19960206	US 1993-141893	19931026
US 6063621	A	20000516	US 1995-407207	19950320

PRIORITY APPLN. INFO.:
US 1992-966923 B2 19921027
US 1993-29340 B2 19930308
US 1993-141893 A2 19931026
US 1995-407207 A2 19950320

AB A method to confer drug resistance on drug-sensitive mammalian cells comprises expression of MRP-encoding nucleic acid in said cells. MRP protects from anthracyclines, epipodophyllotoxins, and Vinca alkaloids. MRP, which belongs to the ABC transporter family, is overexpressed in multidrug resistant cells independently of overexpression of P-glycoprotein. CDNAs encoding two novel human MRPs and a murine MRP were cloned and sequenced. The two human MRP isoform cDNAs differ by only 3 base pairs: T.fwdarw.C at position 2249, C.fwdarw.G at position 4039 and G.fwdarw.C at position 4040 (the proteins differ by Leu685.fwdarw.Ser and Arg1282.fwdarw.Ala). Human MRP mRNA was subject to alternative splicing. MRP was found to be expressed at relatively high levels in lung, testis, and peripheral blood mononuclear cells. Unlike genes for other members of the ABC transporter family, the gene for MRP was found not on chromosome 7, but on chromosome 16. Expression of MRP in drug-sensitive HeLa cells converted them to multidrug-resistance. Both poly- and monoclonal antibodies to MRP were prepd.

IT **179046-57-2P**
RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)
(amino acid sequence; expression of multidrug resistance-assocd.

protein MRP nucleic acid in cells to confer drug resistance)
REFERENCE COUNT: 23 THERE ARE 23 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 11 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:193830 HCAPLUS
DOCUMENT NUMBER: 130:222107
TITLE: Antibodies to multidrug resistant protein MRP and immunoassays for identifying multidrug-resistant tumor cells
INVENTOR(S): Deeley, Roger G.; Cole, Susan P. C.
PATENT ASSIGNEE(S): Queen's University At Kingston, Can.
SOURCE: U.S., 78 pp., Cont.-in-part of U.S. Ser. No. 407,207. CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 8
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5882875	A	19990316	US 1995-462109	19950605
US 5489519	A	19960206	US 1993-141893	19931026
US 6063621	A	20000516	US 1995-407207	19950320

PRIORITY APPLN. INFO.: US 1992-966923 B2 19921027
US 1993-29340 B2 19930308
US 1993-141893 A2 19931026
US 1995-407207 A2 19950320

AB Labeled antibodies or antibody fragments which bind to human MRP and diagnostic kits for identification of multidrug-resistant tumor cells are disclosed. MRP is overexpressed in multidrug resistant cells independently of overexpression of P-glycoprotein. CDNAs encoding two novel human MRPs and a murine MRP were cloned and sequenced. The two human MRP isoform cDNAs differ by only 3 base pairs: T.fwdarw.C at position 2249, C.fwdarw.G at position 4039 and G.fwdarw.C at position 4040 (the proteins differ by Leu685.fwdarw.Ser and Arg1282.fwdarw.Ala). Human MRP mRNA was subject to alternative splicing. MRP was found to be expressed at relatively high levels in lung, testis, and peripheral blood mononuclear cells. Unlike genes for other members of the ABC transporter family, the gene for MRP was found not on chromosome 7, but on chromosome 16. Expression of MRP in drug-sensitive HeLa cells converted them to multidrug-resistance.

IT **179046-57-2P**, Protein MRP (mouse testis multidrug resistance-associated)
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(amino acid sequence; antibodies to multidrug resistant protein MRP and immunoassays for identifying multidrug-resistant tumor cells)
REFERENCE COUNT: 25 THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 12 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:435704 HCAPLUS
DOCUMENT NUMBER: 129:64081
TITLE: Cloning of nucleic acid molecules encoding human and murine multidrug resistance proteins and their diagnostic and therapeutic uses
INVENTOR(S): Deeley, Roger G.; Cole, Susan P. C.
PATENT ASSIGNEE(S): Queen's University at Kingston, Can.
SOURCE: U.S., 82 pp., Cont.-in-part of U. S. Ser. No. 407,207. CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English

FAMILY ACC. NUM. COUNT: 8
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5766880	A	19980616	US 1995-463092	19950605
US 5489519	A	19960206	US 1993-141893	19931026
US 6063621	A	20000516	US 1995-407207	19950320

PRIORITY APPLN. INFO.: US 1992-966923 B2 19921027
US 1993-29340 B2 19930308
US 1993-141893 A2 19931026
US 1995-407207 A2 19950320

AB A novel protein assocd. with multidrug resistance in living cells and capable of conferring multidrug resistance on a cell is disclosed. The protein is assocd. with multidrug resistance which is overexpressed in multidrug resistant cells independently of overexpression of P-glycoprotein. Nucleic acids encoding two novel human multidrug resistance proteins (MRP) and a murine MRP are also disclosed. The two human MRP isoform cDNAs differ by only 3 base pairs: T.fwdarw.C at position 2249, C.fwdarw.G at position 4039 and G.fwdarw.C at position 4040 (the proteins differ by Leu685.fwdarw.Ser and Arg1282.fwdarw.Ala). Transformant cell lines which express the nucleic acid encoding the novel protein are also disclosed. Antibodies which bind the novel multidrug resistance protein are also disclosed. Diagnostic and treatment methods using the novel proteins, nucleic acids, antibodies and cell lines of the invention are also encompassed by the invention.

IT **179046-57-2P**, Protein MRP (mouse testis multidrug resistance-associated)
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(amino acid sequence; cloning of nucleic acid mols. encoding human and murine multidrug resistance proteins and their diagnostic and therapeutic uses)

REFERENCE COUNT: 52 THERE ARE 52 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 13 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:301425 HCAPLUS
DOCUMENT NUMBER: 128:317880
TITLE: Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3
AUTHOR(S): Kawarabayasi, Yutaka; Sawada, Mituhiro; Horikawa, Hiroshi; Haikawa, Yuji; Hino, Yumi; Yamamoto, Saori; Sekine, Mitsuo; Baba, Sin-Ichi; Kosugi, Hiroki; Hosoyama, Akira; Nagai, Yoshimi; Sakai, Mari; Ogura, Keiko; Otsuka, Rie; Nakazawa, Hidekazu; Takamiya, Minako; Ohfuku, Yuhko; Funahashi, Tomomichi; Tanaka, Toshihiro; Kudoh, Yutaka; Yamazaki, Jun; Kushida, Norihiro; Oguchi, Akio; Aoki, Ken-Ichi; Yoshizawa, Takio; Nakamura, Yoshinobu; Robb, Frank T.; Horikoshi, Koki; Masuchi, Yaeko; Shizuya, Hiroaki; Kikuchi, Hisasi
CORPORATE SOURCE: National Institute of Technology and Evaluation, Shibuya, Tokyo, 151-0066, Japan
SOURCE: DNA Research (1998), 5(2), 55-76
CODEN: DARSE8; ISSN: 1340-2838
PUBLISHER: Kazusa DNA Research Institute
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The complete sequence of the genome of a hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3, was detd. by assembling the sequences of the phys. map-based contigs of fosmid clones and of long PCR

products which were used for gap-filling. The entire length of the genome was 1,738,505 bp. The authenticity of the entire genome sequence was supported by restriction anal. of long PCR products, which were directly amplified from the genomic DNA. As the potential protein-coding regions, a total of 2061 open reading frames (ORFs) were assigned, and by similarity search against public databases, 406 (19.7%) were related to genes with putative function and 453 (22.0%) to the sequences registered but with unknown function. The remaining 1202 ORFs (58.3%) did not show any significant similarity to the sequences in the databases. Sequence comparison among the assigned ORFs in the genome provided evidence that a considerable no. of ORFs were generated by sequence duplication. By similarity search, 11 ORFs were assumed to contain the intein elements. The RNA genes identified were a single 16S-23S rRNA operon, two 5S rRNA genes and 46 tRNA genes including 2 with the intron structure. All the assigned ORFs and RNA coding regions occupied 91.25% of the whole genome. The nucleotide and deduced amino acid sequences are available in GenBank Accession Nos. AB009465-AB009531.

IT **207005-80-9**
RL: PRP (Properties)
(amino acid sequence; complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3)

L9 ANSWER 14 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1998:41536 HCAPLUS
DOCUMENT NUMBER: 128:164822
TITLE: P-type ATPases mediate sodium and potassium effluxes in Schwanniomyces occidentalis
AUTHOR(S): Banuelos, Maria A.; Rodriguez-Navarro, Alonso
CORPORATE SOURCE: Departamento de Biotecnologia, Escuela Tecnica Superior de Ingenieros Agronomos, Universidad Politecnica de Madrid, Madrid, 28040, Spain
SOURCE: Journal of Biological Chemistry (1998), 273(3), 1640-1646
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Two genes isolated from Schwanniomyces occidentalis, ENA1 and ENA2, encode P-type ATPases highly homologous to the Na-ATPases of Saccharomyces cerevisiae and complement the Na+ sensitivity of an S. cerevisiae mutant strain lacking its own Na-ATPases. The expression of both ENA1 and ENA2 was highly dependent on a high external pH, but whereas a high pH was sufficient for the expression of ENA2, the expression of ENA1 required a high pH and the presence of Na+. Disruption of ENA1 rendered the cells less tolerant to Na+ than the wild-type strain and decreased their capacity for Na+ extrusion. Disruption of ENA2 did not affect Na+ tolerance, but decreased both the growth at high pH and K+ efflux. We discuss these results and propose that fungal Na-ATPases should be considered alkali cation ATPases. By sequence comparison, we found that fungal Na-ATPases form a homogeneous group that can be distinguished from other cation-pumping P-type ATPases, except from the cta3 Ca-ATPase of Schizosaccharomyces pombe.

IT **202938-24-7**
RL: PRP (Properties)
(amino acid sequence; P-type ATPases mediate sodium and potassium effluxes in Schwanniomyces occidentalis)

REFERENCE COUNT: 44 THERE ARE 44 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 15 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:682849 HCAPLUS

DOCUMENT NUMBER: 127:343084
TITLE: Structure and in vitro substrate specificity of the murine multidrug resistance-associated protein. [Retraction of document cited in CA125:268453]
AUTHOR(S): Paul, Saptarshi; Belinksy, Martin G.; Shen, Hongxie; Kruh, Gary D.
CORPORATE SOURCE: Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, PA, 19111, USA
SOURCE: Biochemistry (1997), 36(45), 13972
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Due to the uncertain validity of the data and the failure to reproduce some of the results in related expts. with membrane vesicles prepd. from another MRP-overexpressing cell line, the article is retracted. The data for Figures 1-3, concerning the structure and expression pattern of murine MRP, are valid.
IT **179046-57-2**
RL: PRP (Properties)
(amino acid sequence; structure and in vitro substrate specificity of murine multidrug resistance-assocd. protein (Retraction))

L9 ANSWER 16 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:652931 HCAPLUS
DOCUMENT NUMBER: 128:2242
TITLE: Synaptopodin: an actin-associated protein in telencephalic dendrites and renal podocytes
AUTHOR(S): Mundel, Peter; Heid, Hans W.; Mundel, Thomas M.; Kruger, Meike; Reiser, Jochen; Kriz, Wilhelm
CORPORATE SOURCE: Department of Anatomy and Cell Biology, University of Heidelberg, Heidelberg, D-69120, Germany
SOURCE: Journal of Cell Biology (1997), 139(1), 193-204
CODEN: JCLBA3; ISSN: 0021-9525
PUBLISHER: Rockefeller University Press
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Synaptopodin is an actin-assocd. protein of differentiated podocytes that also occurs as part of the actin cytoskeleton of postsynaptic densities (PSD) and assocd. dendritic spines in a subpopulation of exclusively telencephalic synapses. Amino acid sequences detd. in purified rat kidney and forebrain synaptopodin and derived from human and mouse brain cDNA clones show no significant homol. to any known protein. In particular, synaptopodin does not contain functional domains found in receptor-clustering PSD proteins. The open reading frame of synaptopodin encodes a polypeptide with a calcd. Mr of 73.7 kDa (human)/74.0 kDa (mouse) and an isoelec. point of 9.38 (human)/9.27 (mouse). Synaptopodin contains a high amt. of proline (.apprx.20%) equally distributed along the protein, thus virtually excluding the formation of any globular domain. Sequence comparison between human and mouse synaptopodin revealed 84% identity at the protein level. In both brain and kidney, in vivo and in vitro, synaptopodin gene expression is differentiation dependent. During postnatal maturation of rat brain, synaptopodin is first detected by Western blot anal. at day 15 and reaches max. expression in the adult animal. The exclusive synaptopodin synthesis in the telencephalon has been confirmed by in situ hybridization, where synaptopodin mRNA is only found in perikarya of the olfactory bulb, cerebral cortex, striatum, and hippocampus, i.e., the expression is restricted to areas of high synaptic plasticity. From these results and expts. with cultured cells the authors conclude that synaptopodin represents a novel kind of proline-rich, actin-assocd. protein that may play a role in modulating actin-based shape and motility of dendritic spines and podocyte foot processes.
IT **198229-20-8**

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; cDNA sequence of human actin-assocd. protein synaptopodin, its specific expression in telencephalon and kidney, and role in neuron and podocyte differentiation)

L9 ANSWER 17 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:408965 HCAPLUS
DOCUMENT NUMBER: 127:91893
TITLE: Sequence analysis of the chlamydomonas reinhardtii flagellar .alpha. dynein gene
AUTHOR(S): Mitchell, David R.; Brown, Kimberly S.
CORPORATE SOURCE: Department of Anatomy and Cell Biology, SUNY Health Science Center, Syracuse, NY, 13210, USA
SOURCE: Cell Motility and the Cytoskeleton (1997), 37(2), 120-126
CODEN: CMCYEO; ISSN: 0886-1544
PUBLISHER: Wiley-Liss
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Flagellar outer row dynein ATPases have been used extensively as model systems for studies of microtubule-based motility. Previously full-length sequences were only available for two of the three catalytic heavy-chain subunits (DHCs) of this enzyme. The authors have completed the sequence of an 18-kb genomic region encoding the Chlamydomonas reinhardtii flagellar outer row dynein .alpha. heavy chain. Unlike the .beta.- and .gamma.-subunits, DHC .alpha. is not required for assembly of other outer row dynein proteins, except for a tightly assocd. light chain, and thus occupies a unique position within this enzyme complex. The predicted 4,499 residue protein retains sequence homol. to other dynein heavy chains throughout its central and C-terminal regions but lacks homol. to any other dyneins in the first 1,000 amino acids, which may account for its unusual assembly properties. This N-terminal domain of DHC .alpha. contains a repetitive sequence rich in alanines, prolines, and glutamic acids. Within the more homologous C-terminal region, which includes the catalytic domain, three short sequences unique to DHC .alpha. may account for its specific catalytic properties and in vivo phosphorylation pattern.
IT **192141-11-0**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; sequence anal. of chlamydomonas reinhardtii flagellar .alpha. dynein gene)

L9 ANSWER 18 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1997:397523 HCAPLUS
DOCUMENT NUMBER: 127:134461
TITLE: Human IgGFc binding protein (Fc.gamma.BP) in colonic epithelial cells exhibits mucin-like structure
AUTHOR(S): Harada, Naoki; Iijima, Shigeyuki; Kobayashi, Kensuke; Yoshida, Takeshi; Brown, William R.; Hibi, Toshifumi; Oshima, Akihiro; Morikawa, Minoru
CORPORATE SOURCE: Tokyo Institute Immunopharmacology, Inc., Chugai Pharmaceutical Co. Ltd., Tokyo, 171, Japan
SOURCE: Journal of Biological Chemistry (1997), 272(24), 15232-15241
CODEN: JBCHA3; ISSN: 0021-9258
PUBLISHER: American Society for Biochemistry and Molecular Biology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Cloning a cDNA for human IgGFc binding protein (Fc.gamma.BP) from human colonic epithelial cells reveals an mRNA and coding region of 17 and 16.2 kilobases, resp. The predicted amino acid sequence contains 12

occurrences of a 400-amino acid cysteine-rich unit resembling that found in mucin. A motif (CGLCGN) in Fc.gamma.BP is conserved in MUC2 and prepro-von Willebrand factor. The N-terminal 450-amino acid sequences are necessary and sufficient to confer IgG Fc binding activity. Fc.gamma.BP mRNA is expressed only in placental and colonic epithelial cells. These results suggest that Fc.gamma.BP may play an important role in immune protection and inflammation in the intestines of primates.

IT **172253-04-2**
RL: PRP (Properties)
(amino acid sequence; sequence and mucin-like structure of human IgGFc binding protein (Fc.gamma.BP) in colonic epithelial cells in relation to tissue expression and IgG Fc binding activity)

L9 ANSWER 19 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:723081 HCAPLUS
DOCUMENT NUMBER: 126:3392
TITLE: A systematic search of the data bases for sequences homologous to titin/connectin
AUTHOR(S): Kolmerer, Bernhard; Olivieri, Nicoletta; Herrmann, Bernhard; Labeit, Siegfried
CORPORATE SOURCE: EMBL Heidelberg, Heidelberg, 69012, Germany
SOURCE: Advances in Biophysics (1996), 33(Muscle Elastic Proteins), 3-11
CODEN: ADVBAT; ISSN: 0065-227X
PUBLISHER: Japan Scientific Societies Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Homologous proteins to Titin/connectins described in Genbank X90568 and X90569 are identified including homol. to ESTs and nebulin and calcium-binding proteins.

IT **171886-19-4**
RL: PRP (Properties)
(systematic search of data bases for sequences homologous to titin/connectin)

L9 ANSWER 20 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:593997 HCAPLUS
DOCUMENT NUMBER: 125:268453
TITLE: Structure and in vitro substrate specificity of the murine multidrug resistance-associated protein
AUTHOR(S): Paul, Saptarshi; Belinsky, Martin G.; Shen, Hongxie; Kruh, Gary D.
CORPORATE SOURCE: Department of Medical Oncology, Fox Chase Cancer Center, Philadelphia, PA, 19111, USA
SOURCE: Biochemistry (1996), 35(42), 13647-13655
CODEN: BICHAW; ISSN: 0006-2960
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Multidrug resistance-assocd. protein (MRP) is a recently described ATP cassette transporter that confers cellular resistance to natural product cytotoxic drugs. To examine the biochem. activity and cellular physiol. of this transporter, we isolated the murine MRP homolog and analyzed its in vitro substrate specificity. Murine MRP transcript is widely expressed in tissues and encodes a protein of 1528 amino acids that is 88% identical to its human homolog. Hydropathy anal. indicated that murine and human MRP, the yeast cadmium resistance transporter, and the sulfonylurea receptor share a conserved topol. distinguished from P-glycoprotein and the cystic fibrosis conductance regulator by an N-terminal hydrophobic region that contains several potential transmembrane domains. Drug uptake assays performed with membrane vesicles prepd. from NIH3T3 cells transfected with a murine MRP expression vector revealed ATP-dependent transport for the natural product cytotoxic drugs daunorubicin and

vincristine, as well as for the glutathione S-conjugates leukotriene C4 and azidophenacyl-S-glutathione (APA-SG). Drug transport was osmotically sensitive and saturable with regard to drug and ATP concns., with Km values of 19 .mu.M, 19 .mu.M, 26 nM, 17 .mu.M, and 77 .mu.M for daunorubicin, vincristine, leukotriene C4, APA-SG, and ATP, resp. Consistent with broad substrate specificity, the drug glutathione conjugate APA-SG, oxidized glutathione, the LTD4 antagonist MK571, arsenate, and genistein were competitive inhibitors of daunorubicin transport, with Ki values of 32 .mu.M, 25 .mu.M, 1.9 .mu.M, 108 .mu.M, and 23 .mu.M, resp. This study demonstrates that the substrate specificity of murine MRP is quite broad and includes both the neutral or mildly cationic natural product cytotoxic drugs and the anionic products of glutathione conjugation. The widespread expression pattern of murine MRP in tissues, combined with its ability to transport both lipophilic xenobiotics and the products of phase II detoxification, indicates that it represents a widespread and versatile cellular defense mechanism.

IT **179046-57-2**
RL: PRP (Properties)
(amino acid sequence; structure and in vitro substrate specificity of murine multidrug resistance-assocd. protein)

L9 ANSWER 21 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1996:368999 HCAPLUS
DOCUMENT NUMBER: 125:108069
TITLE: Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter
AUTHOR(S): Stride, Brenda D.; Valdimarsson, Gunnar; Gerlach, James H.; Wilson, Gerald M.; Cole, Susan P. C.; Deeley, Roger G.
CORPORATE SOURCE: Dep. of Biochemistry, Queen's Univ., Kingston, K7L 3N6, Can.
SOURCE: Molecular Pharmacology (1996), 49(6), 962-971
CODEN: MOPMA3; ISSN: 0026-895X
PUBLISHER: Williams & Wilkins
DOCUMENT TYPE: Journal
LANGUAGE: English

AB In vitro, overexpression of the human multidrug-resistance protein (MRP) causes a form of multidrug resistance similar to that conferred by P-glycoprotein, although the two proteins are only very distantly related. Studies with MRP-enriched membrane vesicles have demonstrated that the protein can bind and transport cysteinyl leukotrienes, as well as some other glutathione conjugates, with high affinity. In contrast, there is no direct evidence of the ability of MRP to bind to transport unmodified forms of the drugs to which it confers resistance. To facilitate studies of the physiol. functions(s) of MRP and its ability to cause multidrug resistance in vivo, we cloned and characterized the mRNA specifying its murine homolog. The murine MRP mRNA encodes a protein of 1528 amino acids that is 88% identical to human MRP. Although detectable by Northern blotting at variable levels in a wide range of tissues, in situ hybridization expts. revealed that MRP mRNA expression in some tissues is cell-type specific. High levels of the mRNA were detected in epithelia lining bronchi and bronchioles, as well as stage-specific expression in the seminiferous epithelium of these testes. Comparison of the predicted hydropathy profiles of human and murine MRP suggests a highly conserved membrane topol., the most distinctive feature of which is an extremely hydrophobic NH2-terminal region contg. five or six potential transmembrane sequences. This structural feature is shared with the sulfonylurea receptor and the yeast cadmium factor 1 but is not present in members of the superfamily, such as the cystic fibrosis transmembrane conductance regulator and P-glycoproteins. Finally, we used overlapping cDNAs to construct an episomally replicating murine MRP expression vector that was stably transfected into HeLa cells. MRP-transfected cell populations

expressed markedly elevated levels of a 180-190 -kDa protein that cross-reacted with a polyclonal antiserum raised against a peptide that is completely conserved in murine and human MRPs. The MRP transfectants also displayed increased resistance to vincristine (5-6-fold) and doxorubicin (<2-fold).

IT **179046-57-2**
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(structure and expression of the mRNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter)

L9 ANSWER 22 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1995:950132 HCAPLUS
DOCUMENT NUMBER: 124:23070
TITLE: The complete nucleotide sequence and genome organization of bean common mosaic virus (NL3 strain)
AUTHOR(S): Fang, G. W.; Allison, R. F.; Zambolim, E. M.; Maxwell, D. P.; Gilbertson, R. L.
CORPORATE SOURCE: Department of Botany and Plant Pathology, Michigan State University, East Lansing, MI, 48824, USA
SOURCE: Virus Research (1995), 39(1), 13-23
CODEN: VIREDF; ISSN: 0168-1702
PUBLISHER: Elsevier
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The nucleotide sequences of 3 cDNA clones corresponding to entire RNA genome of bean common mosaic virus NL3 strain have been detd. The RNA is 9612 nucleotides long, excluding a 3'-terminal poly(A) tail. A putative start codon located at nucleotide positions 170-172 initiates one large open reading frame that is terminated with a UAA codon at position 9368-9370. The predicted polyprotein has 3066 amino acids and an Mr of 340.3 kDa. The positions of putative protein cleavage sites have been detd. by analogy to consensus sequences in other potyviruses. The nucleotide sequences of the non-translated regions and the predicted amino acid sequences of BCMV NL3 were compared with those of other potyviruses. Comparison of the BCMV NL3 proteins with those of other potyviruses indicated a similar genomic organization, and high percentage of amino acid sequence identity in the cylindrical inclusion protein, nuclear inclusion 'b' protein and coat protein. BCMV NL3 displays the highest amino acid sequence identity with soybean mosaic virus.

IT **171760-49-9**
RL: PRP (Properties)
(amino acid sequence; complete nucleotide sequence and genome organization of bean common mosaic virus (NL3 strain))

L9 ANSWER 23 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1995:873218 HCAPLUS
DOCUMENT NUMBER: 124:48553
TITLE: Titins: giant proteins in charge of muscle ultrastructure and elasticity
AUTHOR(S): Labeit, Siegfried; Kolmerer, Bernhard
CORPORATE SOURCE: European Molecular Biology Laboratory, Heidelberg, 69012, Germany
SOURCE: Science (Washington, D. C.) (1995), 270(5234), 293-6
CODEN: SCIEAS; ISSN: 0036-8075
PUBLISHER: American Association for the Advancement of Science
DOCUMENT TYPE: Journal
LANGUAGE: English

AB In addn. to thick and thin filaments, vertebrate striated muscle contains a third filament system formed by the giant protein titin. Single titin mols. extend from Z disks to M lines and are longer than 1 .mu.m. The titin filament contributes to muscle assembly and resting tension, but more details are now known because of the large size of the protein. The

complete complementary DNA sequence of human cardiac titin was detd. The 82-kilobase complementary DNA predicts a 3-megadalton protein composed of 244 copies of Ig and fibronectin type III (FN3) domains. The architecture of sequences in the A band region of titin suggests why thick filament structure is conserved among vertebrates. In the I band region, comparison of titin sequences from muscles of different passive tension identifies two elements that correlate with tissue stiffness. This suggests that titin may act as two springs in series. The differential expression of the springs provides a mol. explanation for the diversity of sarcomere length and resting tension in vertebrate striated muscles.

IT **171886-19-4**
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; characterization of human titins, giant proteins in charge of muscle ultrastructure and elasticity)

L9 ANSWER 24 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1994:647538 HCAPLUS
DOCUMENT NUMBER: 121:247538
TITLE: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes
AUTHOR(S): Mitchell, David R.; Brown, Kimberly S.
CORPORATE SOURCE: Dep. Anat. Cell Biol. Pro., Suny Health Sci. Cent., Syracuse, NY, 13210, USA
SOURCE: Journal of Cell Science (1994), 107(3), 635-44
CODEN: JNCSAI; ISSN: 0021-9533
DOCUMENT TYPE: Journal
LANGUAGE: English

AB We have sequenced genomic clones spanning the complete coding region of one heavy chain (beta) and the catalytic domain of a second (alpha) of the Chlamydomonas reinhardtii flagellar outer arm dynein ATPase. The beta heavy chain gene (ODA-4 locus) spans 20 kb, is divided into at least 30 exons, and encodes a predicted 520 kDa protein. Comparison with sea urchin beta dynein sequences reveals homol. that extends throughout both proteins. Over the most conserved central catalytic region, the Chlamydomonas alpha and beta chains are equally divergent from the sea urchin beta chain (64% and 65% similarity, resp.), whereas the Chlamydomonas gamma chain is more divergent from urchin beta (54% similarity). The four glycine-rich loops identified as potential nucleotide-binding sites in other dynein heavy chains are also present in Chlamydomonas alpha and beta dyneins. Two of these four nucleotide-binding motifs are highly conserved among flagellar dyneins, but only the motif previously identified as the catalytic site in sea urchin dynein is highly conserved between flagellar and cytoplasmic dynein heavy chains. Predictions of secondary structure suggest that all dynein heavy chains possess three large domains, with the four nucleotide-binding consensus sequences located in a central 185 kDa domain that is bounded on both sides by regions that form multiple, short alpha-helical coiled-coils.

IT **158650-91-0**
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(amino acid sequence; sequence anal. of the Chlamydomonas alpha dynein heavy chain (central catalytic region) gene)

L9 ANSWER 25 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1993:487631 HCAPLUS
DOCUMENT NUMBER: 119:87631
TITLE: Complete nucleotide sequences of two soybean mosaic virus strains differentiated by response of soybean containing the Rsv resistance gene
AUTHOR(S): Jayaram, C.; Hill, John H.; Miller, W. Allen
CORPORATE SOURCE: Dep. Plant Pathol., Iowa State Univ., Ames, IA, 50011,

USA
SOURCE: Journal of General Virology (1992), 73(8), 2067-77
CODEN: JGVIAY; ISSN: 0022-1317
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The complete nucleotide sequence of the genomic RNAs of strains G2 and G7 of soybean mosaic virus were detd. In both cases, the genome is 9588 nucleotides long, excluding the 3'-terminal poly(A) sequence. A large open reading frame (nucleotides 132 to 9329) encodes a polyprotein of 3066 amino acids with a predicted Mr of either 349,542 (strain G2) or 349,741 (strain G7). Based on comparison with the proposed locations of cleavage sites of other potyvirus polyproteins, 9 mature proteins are predicted. The mature proteins of the 2 strains share 94-100% amino acid identity, with the greatest variability occurring in the 35K and 42K proteins. Differences in local net charge in portions of these proteins as well as differences in amino acid sequence throughout the genome are discussed in relation to resistance and susceptibility of host plants to strains G2 and G7. Comparison with other potyviruses may be useful for taxonomic clarification of viruses and strains.

IT **149289-62-3 149289-63-4**
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence of, complete)

L9 ANSWER 26 OF 26 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1990:566476 HCAPLUS
DOCUMENT NUMBER: 113:166476
TITLE: TEC1, a gene involved in the activation of Ty1 and Ty1-mediated gene expression in Saccharomyces cerevisiae: cloning and molecular analysis
AUTHOR(S): Laloux, Isabelle; Dubois, Evelyne; Dewerchin, Marianne; Jacobs, Eric
CORPORATE SOURCE: Fac. Sci., Univ. Lib. Bruxelles, Brussels, 1070, Belg.
SOURCE: Molecular and Cellular Biology (1990), 10(7), 3541-50
CODEN: MCEBD4; ISSN: 0270-7306
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Ty and Ty-mediated gene expression obsd. in haploid cells of S. cerevisiae depends on several determinants, some of which are required for the expression of haploid-specific genes. TEC1 Encodes a 486-amino-acid protein that is a trans-acting factor required for full Ty1 expression and Ty1-mediated gene activation. However, mutation or deletion of the TEC1 gene had little effect on total Ty2 transcript levels. TEC1 Is not involved in mating or sporulation processes. Unlike most of the proteins involved in Ty and adjacent gene expression, the product of TEC1 has no known cellular function. Although there was no mating-type effect on TEC1 expression, these results indicate that the TEC1 and the a/.alpha. diploid controls on Ty1 expression are probably not cumulative.

IT **129876-88-6**, Ribonucleic acid formation factor (Saccharomyces cerevisiae clone pILDN486 gene TEC1 reduced)
RL: PRP (Properties)
(amino acid sequence of)

=> select hit rn l9 1-26
E1 THROUGH E19 ASSIGNED

=> fil reg
FILE 'REGISTRY' ENTERED AT 16:52:38 ON 04 JUN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 3 JUN 2003 HIGHEST RN 524916-37-8
DICTIONARY FILE UPDATES: 3 JUN 2003 HIGHEST RN 524916-37-8

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

```
=> s e1-e19
             1 179046-57-2/BI
               (179046-57-2/RN)
             1 171886-19-4/BI
               (171886-19-4/RN)
             1 129876-88-6/BI
               (129876-88-6/RN)
             1 149289-62-3/BI
               (149289-62-3/RN)
             1 149289-63-4/BI
               (149289-63-4/RN)
             1 158650-91-0/BI
               (158650-91-0/RN)
             1 171760-49-9/BI
               (171760-49-9/RN)
             1 172253-04-2/BI
               (172253-04-2/RN)
             1 192141-11-0/BI
               (192141-11-0/RN)
             1 198229-20-8/BI
               (198229-20-8/RN)
             1 202938-24-7/BI
               (202938-24-7/RN)
             1 207005-80-9/BI
               (207005-80-9/RN)
             1 244204-35-1/BI
               (244204-35-1/RN)
             1 253424-06-5/BI
               (253424-06-5/RN)
             1 262986-25-4/BI
               (262986-25-4/RN)
             1 290391-05-8/BI
               (290391-05-8/RN)
             1 300595-83-9/BI
               (300595-83-9/RN)
             1 306331-52-2/BI
               (306331-52-2/RN)
             1 324098-98-8/BI
               (324098-98-8/RN)
L10         19 (179046-57-2/BI OR 171886-19-4/BI OR 129876-88-6/BI OR 149289-62
               -3/BI OR 149289-63-4/BI OR 158650-91-0/BI OR 171760-49-9/BI OR
               172253-04-2/BI OR 192141-11-0/BI OR 198229-20-8/BI OR 202938-24-
               7/BI OR 207005-80-9/BI OR 244204-35-1/BI OR 253424-06-5/BI OR
               262986-25-4/BI OR 290391-05-8/BI OR 300595-83-9/BI OR 306331-52-
               2/BI OR 324098-98-8/BI)
```

```
=>
=>

=>
=> d his l11

     (FILE 'REGISTRY' ENTERED AT 16:52:38 ON 04 JUN 2003)
L11          19 S L10 AND L6

=> d .seq 1-19

L11  ANSWER 1 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   324098-98-8  REGISTRY
CN   Translation initiation factor IF-2; 74568-78972 (Arabidopsis thaliana
     clone F28O16 gene F28O16.19) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AC010718-derived protein GI 6143897
SQL  1146
RN   324098-98-8  REGISTRY

SEQ     51 KVVITGKKKG KKRNKKGTQQ QQDDDDDFAD KFAVEEEVVP DNAFVGNNKK
                                                    =====
       501 EAQHVVKKEF KAHYSDHETE KPTAKPAGMS KLETAAVKAI SEVEDAATQT
                       ===
HITS AT:   86-90, 513-515

REFERENCE   1:  134:173731

L11  ANSWER 2 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   306331-52-2  REGISTRY
CN   Titin (Drosophila melanogaster) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AJ271740-derived protein GI 8250181
SQL  16215
RN   306331-52-2  REGISTRY

SEQ    301 YEISYSSGVA TLRVKNATAR DGGHYTLLAE NLQGCVVSSA VLAVEPAAET
           ===
       701 AESRAILSVV QRPSIEQSSQ NPNSLQYINQ LEDYSRYQRT ESIDEQLNQA
                                             ===
      6351 DVKVVAVSED VLPEEEVVPT EETPEAKQKA HKKRTKRLKE ASVEGQPQLL
                        =====
     14101 DELTVKVEEE VVPEPIVEEE VIEEFEIKKK PKEPEPEDIV DAAIVKLKKP
                   == ===
HITS AT:   304-306, 733-735, 6365-6369, 14109-14113

REFERENCE   1:  133:347455

L11  ANSWER 3 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   300595-83-9  REGISTRY
CN   Titin (human clone #14104 9448-amino acid fragment) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AJ277892-derived protein GI 8249467
CN   Titin (human clone #14104 gene TTN 9448-amino acid fragment)
SQL  9448
RN   300595-83-9  REGISTRY

SEQ   1251 LSESNTVRMY FVNSEAILDI TDVKVEDSGS YSCEAVNDVG SDSCSTEIVI
                                           = ==
      4351 ESCNISLEDF VTELELFEVQ PLESGDYSCL VTNDAGSASC TTHLFVKEPA
                                      ===
      5001 DKGEIVRESD NIWISYSENI ATLQFSRVEP ANAGKYTCQI KNDAGMQECF
```

```
                          ===
      5301 SVAELELFDV DTSQSGEYTC IVSNEAGKAS CTTHLYIKAP AKFVKRLNDY
                                                              ==
      5351 SIEKGKPLIL EGTFTGTPPI SVTWKKNGIN VTPSQRCNIT TTEKSAILEI
           =
      6951 VPQRVEVTRH EVSAEEEWSY SEEEEGVSIS VYREEEREEE EEAEVTEYEV
                               == =
      7401 EYIHEEEEFI TEEEVVPVIP VKVPEVPRKP VPEEKKPVPV PKKKEAPPAK
                            =====
HITS AT:   1280-1282, 4376-4378, 5015-5017, 5349-5351, 6969-6971,
           7413-7417

REFERENCE   1:  133:291704

L11  ANSWER 4 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   290391-05-8  REGISTRY
CN   Iron(III) ABC transporter, permease protein (Vibrio cholerae strain N16961
     gene VC0203) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AE004110-derived protein GI 9654609
SQL  653
RN   290391-05-8  REGISTRY

SEQ    101 IVNIWFSDWV ADYSALAAMA GALLAFALII SIAGLRNLTG LPLVVSGMVV
                     ===
       251 IGFIGLLTPN IARSLGARTT KMELYSSALL GALLLLATDM LAMGLSVWAE
                                                              =
       301 EVVPSGITAA VIGAPALIWF SRRQLQAQDS LSISLSSHRR SPSRWAVMLI
           ====
HITS AT:   112-114, 300-304

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  133:218310

L11  ANSWER 5 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   262986-25-4  REGISTRY
CN   Protein (Drosophila melanogaster gene CG6769) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AE003507-derived protein GI 7293398
SQL  409
RN   262986-25-4  REGISTRY

SEQ    251 TFYSLDAVRK HMIDKGHCQM LHEGVALAEY AEYYDYSSSY PDNNEGMDTD
                                                        ===
       301 EEVVPDLLDG DEYQLVLPSG AVIGHRSLLR YYKQRLRPER AVVIKKSDRK
           =====
HITS AT:   285-287, 301-305

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  132:304167

L11  ANSWER 6 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   253424-06-5  REGISTRY
CN   Protein (human clone hg04224 gene KIAA1170 C-terminal fragment) (9CI)  (CA
     INDEX NAME)
OTHER NAMES:
CN   GenBank AB032996-derived protein GI 6330197
SQL  838
RN   253424-06-5  REGISTRY

SEQ    451 QDTDTLVGLP RPIHESVKTL KQHKYISIAD VQIKNEEELE KCPMSLGEEV
```

```
                                                                    ===
       501 VPETPCEILY QGMLYSLPQY MIALLKILLA AAPTSKAKTD SINILADVLP
           ==
       801 VLGQRLDLPE DFHYSYELWL EREVFSQPIC WEELLQNH
                    ===
HITS AT:    498-502, 813-815

REFERENCE   1:  132:74206

L11  ANSWER 7 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   244204-35-1  REGISTRY
CN   KIAA1029 protein (human clone fh00363 gene KIAA1029) (9CI)  (CA INDEX
     NAME)
OTHER NAMES:
CN   GenBank AB028952-derived protein GI 5689395
SQL  903
RN   244204-35-1  REGISTRY

SEQ    301 GQRSPASERR PLGNFTAPPT YTETLSTAPL ASWVRSPPSY SVLYPSSDPK
                                                 == =
       401 TADEKRRQRD QGEVGVEEEP FALGAEASNF QQEPAPRDRA SPAAAEEVVP
                                                       =====
HITS AT:    339-341, 446-450

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  131:238581

L11  ANSWER 8 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   207005-80-9  REGISTRY
CN   431Aa long protein (Pyrococcus horikoshii gene PHAB028) (9CI)  (CA INDEX
     NAME)
OTHER NAMES:
CN   GenBank AB009484-derived protein GI 3130856
SQL  431
RN   207005-80-9  REGISTRY

SEQ     51 RIEDPYTLIR ALIYIGYLSG LTGLKSARKA FREAISYSEV LPKELRDQII
                                                 ===
       301 EIREKIIEIM EEGDERFSSI VKTITEKTSN EEILIGAVKY FLRLDEFEEV
                                                       ===
       351 VPLLRRIRTE KGKSIALGFI AYHLINKGRI GDAIDIVLEI KDRNLASKLA
           ==
HITS AT:    86-88, 348-352

**RELATED SEQUENCES AVAILABLE WITH SEQLINK**

REFERENCE   1:  128:317880

L11  ANSWER 9 OF 19  REGISTRY  COPYRIGHT 2003 ACS
RN   202938-24-7  REGISTRY
CN   Phosphatase, adenosine tri- (Schwanniomyces occidentalis gene ENA2) (9CI)
     (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AF030861-derived protein GI 2623238
CN   P-type ATPase 2 (Debaryomyces occidentalis gene ENA2)
SQL  1082
RN   202938-24-7  REGISTRY

SEQ    151 NGDDTTIPAE EVVPGDIVHI KVGDTVPADL RLIDLMNLET DEALLTGESL
                   = ====
       201 PITKNHLDVY DDYSVPIPVG DRLNLAYSSS VVSKGRGTGI VIATALDTQI
                     ===
```

651 KSCHNAGINV HMLTGDHPGT AKAIAQEVGI LPHNLYHYSE EVVKAMCMTA
===

HITS AT: 160-164, 212-214, 687-689

REFERENCE 1: 128:164822

L11 ANSWER 10 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **198229-20-8** REGISTRY
CN Synaptopodin (human clone 178792/166347/167192 actin-associated) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank Y11072-derived protein GI 2654323
SQL 685
RN **198229-20-8** REGISTRY

SEQ 301 GQRSPASERR PLGNFTAPPT YTETLSTAPL ASWVRSPPSY SVLYPSSDPK
== =
401 TADEKRRQRD QGEVGVEEEP FALGAEASNF QQEPAPRDRA SPAAAEEVVP
=====

HITS AT: 339-341, 446-450

REFERENCE 1: 128:2242

L11 ANSWER 11 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **192141-11-0** REGISTRY
CN Dynein (Chlamydomonas reinhardtii gene ODA11 .alpha.-chain) (9CI) (CA INDEX NAME)
SQL 4499
RN **192141-11-0** REGISTRY

SEQ 1451 PRFYFVSSAD LLDILSNGNN PMRVQIHMNK CFQAIDNVRL DSEEVVPGRR
=====
2201 CTWLREMFDK YIPPTLLEMK KSYSHITPLA QMNFISTLVN IMEGVLKPEN
===
2701 ISRIVSNPSG HALLVGVGGS GKQSLARLAA HICGYATQMI VISGSYSMNN
===

HITS AT: 1493-1497, 2222-2224, 2745-2747

REFERENCE 1: 127:91893

L11 ANSWER 12 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **179046-57-2** REGISTRY
CN Protein MRP (mouse clone 14B multidrug resistance reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 2: PN: US6001563 SEQID: 2 claimed protein
CN 6: PN: US6025473 SEQID: 6 claimed protein
CN Multidrug resistance protein MRP (mouse clone 14B reduced)
CN Protein MRP (mouse testis multidrug resistance-associated)
SQL 1528
RN **179046-57-2** REGISTRY

SEQ 251 EEVVPVLVNN WKKECDKSRK QPVRIVYAPP KDPSKPKGSS QLDVNEEVEA
=====
1201 ECVGNCIVLF AALFAVISRH SLSAGLVGLS VSYSLQITAY LNWLVRMSSE
===

HITS AT: 251-255, 1232-1234

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 132:176625

REFERENCE 2: 132:19669

REFERENCE 3: 130:247846

REFERENCE 4: 130:222107

REFERENCE 5: 129:64081

REFERENCE 6: 127:343084

REFERENCE 7: 125:268453

REFERENCE 8: 125:108069

L11 ANSWER 13 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **172253-04-2** REGISTRY
CN Receptor, immunoglobulin G Fc chain (human colon) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN IgG Fc binding protein (human colon epithelial cell)
CN IgG Fc binding protein (human)
CN Protein Fc.gamma.BP (IgG Fc region-binding) (human colon epithelial cell)
SQL 5405
RN **172253-04-2** REGISTRY

```
SEQ   1401 FQKPNGSQAG NANEFGNSWE EVVPDSPCLP PTPCPPGSED CIPSHKCPPE
                                 = ====
      2601 DFQKPNGSQA GNANEFGNSW EEVVPDSPCL PPPTCPPGSE GCIPSEECPP
                                 =====
      3801 DDFQKPNGSQ AGNANEFGNS WEEVVPDSPC LPPPTCPPGS EGCIPSEECP
                                  =====
      5101 CQAAGVAVKP WRTDSFCPLH CPAHSHYSIC TRTCQGSCAA LSGLTGCTTR
                                        ===
HITS AT:   1420-1424, 2621-2625, 3822-3826, 5126-5128
```

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 135:370241

REFERENCE 2: 127:134461

REFERENCE 3: 124:137820

L11 ANSWER 14 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **171886-19-4** REGISTRY
CN Connectin (human skeletal muscle isoform N2-A fragment reduced) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Titin (human skeletal muscle isoform N2-A fragment reduced)
SQL 7962
RN **171886-19-4** REGISTRY

```
SEQ    451 MYFVNSEAIL DITDVKVEDS GSYSCEAVND VGSDSCSTEI VIKEPPSFIK
                                  ===
      3551 DFVTELELFE VQPLESGDYS CLVTNDAGSA SCTTHLFVKE PATFVKRLAD
                               ===
      4201 SDNIWISYSE NIATLQFSRV EPANAGKYTC QIKNDAGMQE CFATLSVLEP
              ===
      4501 DVDTSQSGEY TCIVSNEAGK ASCTTHLYIK APAKFVKRLN DYSIEKGKPL
                                                      ===
      6151 RHEVSAEEEW SYSEEEEGVS ISVYREEERE EEEEAEVTEY EVMEEPEEYV
                     ===
      6601 FITEEEVVPV IPVKVPEVPR KPVPEEKKPV PVPKKKEAPP AKVPEVPKKP
            =====
HITS AT:   472-474, 3568-3570, 4207-4209, 4541-4543, 6161-6163,
```

6605-6609

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 126:3392

REFERENCE 2: 124:48553

L11 ANSWER 15 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **171760-49-9** REGISTRY
CN Polypeptide (bean common mosaic virus strain NL3) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank U19287-derived protein
SQL 3062
RN **171760-49-9** REGISTRY

SEQ 501 ALINPSLLCD NQLDRNGNFV WGERGRHSKR FFENFFEEVV PSEGYKKYVI
==== =
2051 DYSSVSTLIC RLVNSSDGHN ETIYGIGYGS YIITNGHLFR RNNGTLTVKT
===
HITS AT: 537-541, 2051-2053

REFERENCE 1: 124:23070

L11 ANSWER 16 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **158650-91-0** REGISTRY
CN Dynein (Chlamydomonas reinhardtii clone A1.2/A1.1/A2.2 gene ODA11 heavy chain fragment reduced) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Dynein (Chlamydomonas reinhardti clone A1.2/A1.1/A2.2 gene ODA11 heavy chain fragment reduced)
OTHER NAMES:
CN Dynein (Chlamydomonas reinhardti clone A1.2/A1.1/A2.2 gene ODA11 heavy chain isoform alpha catalytic region fragment reduced)
SQL 2404
RN **158650-91-0** REGISTRY

SEQ 351 SEEVVPGRRP KALGMESCVG IEYVPFSSLP LENKVEQYMN DIIAKMRNDV
=====
1051 RAKWKDPQLP CTWLREMFDK YIPPTLLEMK KSYSHITPLA QMNFISTLVN
===
1601 VISGSYSMNN FKEDIQKMYK RTKVKGEGVM FLFTDSQIVD ERMLVYINDL
===
HITS AT: 352-356, 1082-1084, 1605-1607

REFERENCE 1: 121:247538

L11 ANSWER 17 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **149289-63-4** REGISTRY
CN Protein, poly- (soybean mosaic virus strain G7 reduced) (9CI) (CA INDEX NAME)
SQL 3065
RN **149289-63-4** REGISTRY

SEQ 301 MEDIQHYSQN PEAQFFRGWK KVFDKMPPHV ENHECTIDFT NEQCGELAAA
===
501 LLCDNQLDKN GNFVWGERGR HSKRFFANYF EEVVPSEGYS KYVIRKNPNG
=====
2001 QNSNAIAGFP EREDELRQTG LPQVVSKSDV PRAKERVEME SKSVYKGLRD
=
2051 YSGISILICQ LTNSSDGHKE TMFGVGYGSF IITNGHLFRR NNGMLTVKTW
==
HITS AT: 306-308, 531-535, 2050-2052

REFERENCE 1: 119:87631

L11 ANSWER 18 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **149289-62-3** REGISTRY
CN Protein, poly- (soybean mosaic virus strain G2 reduced) (9CI) (CA INDEX NAME)
SQL 3065
RN **149289-62-3** REGISTRY

SEQ 301 MEDIQHYSQN PEAQFFRGWK KVFDKMPPHV ENHECTIDFT NEQCGELAAA
===
501 LLCDNQLDKN GNFVWGERGR HSKRFFANYF EEVVPSEGYS KYVIRTNPNG
=====
2001 QNSNAIAGFP EREDELRQTG LPQVVSKSDV PRAKERVEME SKSVYKGLRD
=
2051 YSGISTLICQ LTNSSDGHKE TMFGVGYGSF IITNGHLFRR NNGMLTVKTW
==
HITS AT: 306-308, 531-535, 2050-2052

REFERENCE 1: 119:87631

L11 ANSWER 19 OF 19 REGISTRY COPYRIGHT 2003 ACS
RN **129876-88-6** REGISTRY
CN RNA formation factor (Saccharomyces cerevisiae clone pILDN486 gene TEC1 reduced) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Ribonucleic acid formation factor (Saccharomyces cerevisiae clone pILDN486 gene TEC1 reduced)
SQL 486
RN **129876-88-6** REGISTRY

SEQ 1 MSLKEDDFGK DNSRNIESYT GRIFDVYIQK DSYSQSALDD MFPEAVVSTA
===
351 KKIEIEQRKI INKYQRISRI QEHESNPEFS SNSNSGSEYE SEEEVVPRSA
=====
HITS AT: 32-34, 393-397

RELATED SEQUENCES AVAILABLE WITH SEQLINK

REFERENCE 1: 113:166476